



Krakowska Akademia
im. Andrzeja Frycza
Modrzewskiego

Boris Kožuh

Elena Boczarowa

**BADANIA
METAANALITYCZNE
W PEDAGOGICE**

Kraków 2009

Rada Wydawnicza Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego:
Klemens Budzowski, Maria Kapiszewska, Zbigniew Maciąg, Jacek M. Majchrowski

Recenzja:
prof. nadzw. dr hab. Barbara Sitarska

Projekt okładki:
Joanna Sroka

Korekta
zespół

Copyright© by Krakowska Akademia im. Andrzeja Frycza Modrzewskiego
Kraków 2009

ISBN 978-83-7571-042-7

Żadna część tej publikacji nie może być powielana ani magazynowana w sposób umożliwiający ponowne wykorzystanie, ani też rozpowszechniana w jakiegokolwiek formie za pomocą środków elektronicznych, mechanicznych, kopiujących, nagrywających i innych, bez uprzedniej pisemnej zgody właściciela praw autorskich

Na zlecenie:
Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego
www.ka.edu.pl

Wydawca:
Krakowskie Towarzystwo Edukacyjne sp. z o.o. – Oficyna Wydawnicza AFM,
Kraków 2009

Sprzedaż prowadzi:
Księgarnia Krakowskiego Towarzystwa Edukacyjnego sp. z o.o.
Kampus Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego
ul. Gustawa Herlinga-Grudzińskiego 1
30-705 Kraków
tel./faks: (012) 252 45 93
e-mail: ksiegarnia@kte.pl

Skład:
Oleg Aleksejczuk

Druk i oprawa:

Wstęp

W badaniach edukacji trudno wskazać metodę, która jako jedyna mogłaby posłużyć w badanej sytuacji i jako jedyna mogłaby udzielić odpowiedzi na postawione pytania badawcze. Z technikami badawczymi jest podobnie: potrzebne dane można bowiem zgromadzić różnymi technikami. W przypadku metod statystycznych istnieje ta sama zależność – przy opracowywaniu danych w celu dotarcia do tego samego lub bardzo zbliżonego celu można wybrać różne rozwiązania.

Niniejsza książka ilustruje to samo podobieństwo, pokazuje bowiem, jak prawa metodologiczne pojmują inni autorzy. Porównując te poglądy z tezami głoszonymi przez innych polskich autorów, czytelnik z pewnością odnajdzie różnice. Poniższe opracowanie jest jednak monografią, a nie podręcznikiem metodologii i różnice w podejściach metodologicznych nie są tu zasadnicze. Można nawet zaryzykować twierdzenie, że odmiennosc ta może być dobrym wstępem do zrozumienia metodologii. Studiowanie praw metodologii i ich analiza nie są przecież możliwe jedynie na poziomie podręczników. Zgłębianie tajemnic metodologii badań umożliwiają dopiero monografie. Stąd te trochę inne poglądy i inne interpretacje, a szczególnie odmienne klasyfikacje mogą stać się bodźcem do głębszych przemyśleń teoretycznych.

Głównym zamierzeniem tego opracowania jest ukazanie zależności doświadczeń teoretycznych i empirycznych w rozwoju metodologii pedagogicznej. Autorzy podjęli tu próbę prezentacji i analizy podstawowych teorii, koncepcji i kontrowersji narosłych wokół integracji badań, które wpłynęły na burzliwy rozwój tej subdyscypliny nauk pedagogicznych.

Celem niniejszej książki jest przedstawienie polskim czytelnikom nowego podejścia w badaniach edukacji. Ogrom nowej wiedzy, której dostarcza praktyka badawcza w edukacji, można objąć i zintegrować przede wszystkim za pomocą metaanalizy. Pierwsze dwa rozdziały dotyczą systematyki i budowy metodologii pedagogiki w celu umiejscowienia nowej, prezentowanej tu metody badań, którą jest metaanaliza. W rozdziale trzecim opisano i zdefiniowano podstawowe pojęcia dotyczące integracji i metaanalizy.

Wstęp

Pozostałe rozdziały szczegółowo opisują metaanalizę wraz z prezentacją procedur potrzebnych do przeprowadzania badań metaanalizycznych. Pomimo że metaanaliza nie jest pojęciem nowym, bo jej rozwój przypada na początek lat siedemdziesiątych dwudziestego wieku, to jej przesłanie, choć ciągle słabo znane, jest do dziś niezwykle aktualne i niezbędne do przeprowadzania badań w dziedzinie nauk społecznych. Integracja ta stanowi bowiem pomost pomiędzy zdarzeniami specyficznymi i faktami już rozpoznanymi a prawami o charakterze bardziej ogólnym. Można przyjąć założenie, iż próba początkowej integracji stanowi już całkiem dobre pierwotne badanie empiryczne. Dokonana w ten sposób integracja jest zarówno jakościowa, jak i ilościowa. Historia integracji pokazuje wyraźnie, iż istniały różne tendencje w ustalaniu proporcji zarówno w jakościowej, jak i ilościowej procedurze dociekań naukowych. Przykładem integracji głównie jakościowej są znane od dawna w pedagogice przeglądy tradycyjne.

Do dziś wiele pytań zadanych przez twórców metaanalizy pozostaje bez odpowiedzi. Dlatego nadal istnieje potrzeba poszukiwania nowych konstruktywnych rozwiązań znanych już wątpliwości oraz otwierania kolejnych dróg w integracji badań empirycznych. Efektem tych poszukiwań były publikacje pierwszych teoretycznych prac dotyczących idei, sensu i podstaw prowadzenia integracji badań. Szybko stało się jasne i oczywiste, że do przeprowadzania integracji są potrzebne nowe metody statystyczne. Ich rozwój stał się możliwy głównie dzięki poszukiwaniom oraz rozważaniom Glassa, Smitha, Rosenthala i Hedgesa. Sposoby przez nich wypracowane stanowią narzędzie, dzięki któremu możliwe jest znalezienie odpowiedzi na dwa zasadnicze pytania metaanalizy. Jedno z nich dotyczy określenia mocy efektu, stanowiącej wynik otrzymany z badań pojedynczych, drugie natomiast zawiera problem wpływu poszczególnych cech badań na ich rezultaty.

W opracowaniu autorzy zwrócili też uwagę czytelnika na rodzące się koncepcje integracji ilościowych i jakościowych oraz związane z nimi polemiki. Wśród trwających sporów i kontrowersji niepodważalne staje się dziś ustalenie, że integracja rezultatów badań empirycznych jest obecnie jednym z ważniejszych aspektów badań naukowych.

Czytelnikom życzymy satysfakcji z refleksji nad lekturą oraz odnalezienia odpowiedzi na nurtujące ich problemy badawcze. Zawilości niektórych obszarów metaanalizy, które nie zostały tu ostatecznie rozwiązane lub nawet podjęte z uwagi na konieczność wyeksponowania zasadniczego problemu tej monografii, zostaną z pewnością omówione w kolejnych opracowaniach.

Autorzy

Rozdział I

Podstawy metodologii pedagogiki

1.1. Wstęp

Problemy metodologiczne psychologii i pedagogiki zawsze były zaliczane do najbardziej aktualnych i palących w rozwoju myśli psychologiczno-pedagogicznej. Badanie zjawisk psychologicznych i pedagogicznych z pozycji dialektyki, tj. nauki o najbardziej ogólnych prawach rozwoju przyrody, społeczeństwa i myśli, pozwala ujawnić ich jakościową odrębność, powiązania z innymi zjawiskami i procesami społecznymi.

Zgodnie z zasadami tej teorii, edukację, wychowanie i rozwój przyszłych specjalistów należy rozpatrywać w bezpośrednim związku z konkretnymi warunkami życia społecznego i działalności zawodowej. Wszystkie zjawiska psychologiczno-pedagogiczne badane są z uwzględnieniem ich zmian i rozwoju, ukazania zachodzących sprzeczności i sposobów ich rozwiązywania.

Z filozofii wiemy, że **metodologia jest nauką o najbardziej ogólnych zasadach poznania i przekształcenia obiektywnej rzeczywistości, przebiegu i przejawach tego procesu**. Obecnie rola metodologii w ukazaniu perspektyw rozwoju nauki pedagogicznej istotnie wzrosła.

Po pierwsze, w nowoczesnej nauce widoczne są tendencje zmierzające do integracji wiedzy, kompleksowej analizy różnych zjawisk obiektywnej rzeczywistości. Na przykład, dziś w naukach społecznych szeroko wykorzystywane są osiągnięcia cybernetyki, matematyki, teorii prawdopodobieństwa i innych nauk, których wcześniej nie uwzględniano podczas realizacji funkcji metodologicznych w konkretnym badaniu społecznym. Wyraźnie wzmocniły się związki między samymi naukami i kierunkami naukowymi. W wyniku tych procesów coraz bardziej zacierają się granice między teorią pedagogiczną a ogólnopsychologiczną koncepcją jednostki; między analizą ekonomiczną

problemów społecznych a badaniem psychologiczno-pedagogicznym jednostki; między pedagogiką a genetyką, pedagogiką a fizjologią itd. Warto w tym miejscu podkreślić, że obecnie integracja wszystkich nauk ma jasno wybrany obiekt – człowieka. Właśnie dlatego coraz ważniejszą rolę w połączeniu wysiłków różnych gałęzi nauki skierowanych na badanie człowieka odgrywają psychologia i pedagogika.

Biorąc pod uwagę fakt, że psychologia i pedagogika coraz bardziej wchłaniają osiągnięcia różnych dziedzin nauki, wzmacniają się jakościowo i ilościowo, stale wzbogacając i poszerzając swój zakres, powstaje problem, by ten wzrost był uświadomiony, skorygowany, kierowany, co z kolei bezpośrednio zależy od metodologicznego uświadomienia tego zjawiska. W ten sposób metodologia odgrywa główną rolę w badaniach psychologiczno-pedagogicznych, nadaje im naukową całościowość, systematyczność, ukierunkowanie zawodowe i podnosi efektywność.

Po drugie, stały się bardziej skomplikowane same nauki psychologii i pedagogiki, gdy pojawiło się więcej różnych metod badań, a także rozszerzyły się granice zakresu ich prac badawczych. W tej sytuacji należy, z jednej strony, pamiętać o właściwym przedmiocie badań – rozwiązywaniu problemów psychologiczno-pedagogicznych, a z drugiej – nie pogubić się w ogromie faktów empirycznych i kierować konkretne badania na rozwiązanie podstawowych problemów psychologii i pedagogiki.

Po trzecie, obecnie oczywista stała się dysproporcja między problemami filozoficzno-metodologicznymi a bezpośrednią metodologią badań psychologiczno-pedagogicznych. Z jednej strony mamy tu do czynienia z problemami filozofii psychologii i pedagogiki, a z drugiej – ze specjalistycznymi kwestiami metodologicznymi badań psychologiczno-pedagogicznych. Podsumowując, psychologowie i pedagodzy coraz częściej zderzają się z problemami, które wychodzą poza ramy konkretnego badania, tj. problemami metodologicznymi jeszcze nierozwiązanymi przez nowoczesną filozofię.

Po czwarte, dziś psychologia i pedagogika mają charakter szczególnego pola doświadczalnego zastosowania matematycznych metod w naukach społecznych – potężnym bodźcem rozwoju całych działów matematyki. W tym obiektywnym procesie wzrostu, doskonalenia systemu metodycznego danych nauk nieuchronnie pojawiają się elementy absolutyzacji ilościowych metod badania, które z kolei negatywnie wpływają na analizę jakościową.

Należy jednak pamiętać, że chociażby z przyczyn gnoseologicznych, wykorzystując metody matematyczne, można nie tylko zbliżyć się do prawdy, ale wręcz oddalić się od niej. Żeby uniknąć takiego zagrożenia, analizę ilościową należy uzupełniać analizą jakościową, tj. metodologiczną.

W takim wypadku metodologia pomo e uniknąć popełniania błędu, pozwoli wybrnąć z niezliczonych korelacji, da mo liwość wybrania dla analizy jakościowej najbardziej istotnych zależności statystycznych i wyciągnięcia poprawnych wniosków z ich analizy.

Po piąte, człowiek pozostaje siłą rozstrzygającą w działalności zawodowej. Stwierdzenie to niejako wynika z ogólnosocjologicznego prawa wzrastania roli czynnika subiektywnego w historii, w rozwoju społeczeństwa wynikającym z postępu społecznego.

Z wy ej przytoczonych dociekań mo na wyciągnąć w pełni uzasadnione stwierdzenie, e w obecnych warunkach znaczenie metodologii w badaniach psychologiczno-pedagogicznych stale wzrasta. W tym miejscu trzeba uściślić, co nale y rozumieć przez termin „metodologia”, co jest jej istotą, strukturą i poziomem logicznym, jakie funkcje spełnia.

Termin „**metodologia**” pochodzi z języka greckiego i oznacza „naukę o metodzie” lub „teorię metody”. W nowoczesnej nauce metodologia pojmowana jest w wąskim i szerokim sensie tego wyrazu. W szerszym znaczeniu metodologia oznacza zbiór najbardziej ogólnych, przede wszystkim światopoglądowych zasad słu ących do rozwiązywania złożonych zadań teoretycznych i praktycznych, co składa się na światopoglądową pozycję badacza. Jednocześnie, jest to nauka o metodach poznania, uzasadniająca zasady wyjściowe i sposoby ich konkretnego zastosowania w działalności poznawczej i praktycznej. Metodologię w wąskim znaczeniu tego słowa mo na określić jako naukę o metodach badań naukowych.

Podsumowując, w nowoczesnej literaturze naukowej metodologię najczęściej pojmuje się jako naukę o zasadach budowy, formach i sposobach działalności naukowo-poznawczej. Metodologia nauki daje charakterystykę składników badania naukowego – jego obiektu, przedmiotu, zadań badania, ogółu metod, środków i sposobów badawczych niezbędnych dla ich rozwiązania, a tak e umo liwia spojrzenie na konsekwencje działalności badacza w procesie rozwiązywania problemu naukowego. Wraz z rozwojem nauki, kształtowaniem się jej jako realnej siły wytwórczej, ujawnia się charakter współzależności między działalnością naukową a działalnością praktyczną, która w coraz większym stopniu opiera się na wynikach nauki. Znajduje to odzwierciedlenie w ukazaniu metodologii jako nauki o metodach poznania naukowego, skierowanego na przekształcenie świata.

Nale y również uwzględnić fakt, e wraz z rozwojem nauk społecznych pojawiają się szczegółowe teorie działalności. Na przykład, jedną z takich teorii jest teoria pedagogiczna, w której skład wchodzi szereg szczegółowych teorii wychowania, edukacji, rozwoju, zarządzania systemem oświaty

itd. Prawdopodobnie powyższy fakt doprowadził do jeszcze węższego rozumienia metodologii jako nauki o zasadach, budowie, formach i sposobach działalności naukowo-poznawczej.

Co należy rozumieć przez pojęcie metodologia pedagogiki? Spróbujmy odpowiedzieć na to pytanie bardziej szczegółowo.

W kręgu pojęć, z którymi styka się pedagog, termin „metodologia” występuje jako jeden z najbardziej nieokreślonych, wieloznacznych, a nawet odstrasających. Przyczyn tego faktu należy szukać poza zakresem nauki pedagogicznej. Obecnie metodologia, z jednej strony, jest traktowana zarówno jako system zasad i sposobów budowy działalności teoretycznej i praktycznej, jak i nauka o tym systemie; z drugiej – jako nauka o metodach poznania naukowego i przekształcenia świata; z trzeciej zaś – jako nauka o pozycjach ideowych nauki, logice jej rozwoju i metodach badawczych. Konkretyzując powyższe podejścia wyodrębnia się trzy poziomy metodologii: ogólny, specjalistyczny i szczegółowy.

Metodologia ogólna podaje najbardziej precyzyjne definicje dotyczące powszechnych praw rozwoju świata obiektywnego, o miejscu i roli w nim tych zagadnień, które bada dana nauka.

Metodologia specjalistyczna lub metodologia konkretnej nauki pozwala tej ostatniej formułować własne prawidłowości, odnoszące się do odrębności rozwoju i funkcjonowania badanych przez nią zagadnień.

Metodologia szczegółowa przejawia się jako ogół metod, sposobów i metodyk badania konkretną nauką różnych zjawisk, które są przedmiotem i obiektem jej analizy.

Rozpatrzenie zjawisk pedagogicznych na wszystkich trzech poziomach metodologii ukazuje *pełny obraz naukowy pedagogiki*, tj. następuje zastępowanie nauki jej metodologią lub utożsamianie, przyrównywanie pojęć „nauka” i „metodologia” oraz ich synonimizacja. Takie podejście ewidentnie przekracza zakres metodologii w systemie wiedzy naukowej i wymaga uściślenia jej przedmiotu.

Dokładna analiza wymienionych poziomów metodologii ukazuje, że metodologia ogólna jest niczym innym, jak *podstawą filozoficzną nauki*, w tym przypadku – pedagogiki; metodologia specjalistyczna jest *podstawą teoretyczną* pedagogiki i tylko metodologia szczegółowa odpowiada pierwotnemu przeznaczeniu metodologii jako **nauki o metodzie badań** (analiza terminologiczna). Naukę o metodach badań naukowych można rozpatrywać jako:

- **metodologię powszechną**, będącą światopoglądową pozycją i ogółem metod filozoficznego poznania świata i własnego miejsca w świecie (metoda dialektyczno-materialistyczna, metoda indukcyjno-dedukcyjna, de-

terministyczna [ukazanie zależności przyczynowo-skutkowych], metoda analizy konkretno-historycznej itd.);

- **metodologię ogólnonaukową**, zawierającą podejście systemowe i całościowe, optymalizację wybranych procesów, podejście aktywne i aksjologiczne, charakterystykę różnych typów badań naukowych, ich etapy, elementy itd.;
- **metodologię konkretno-naukową** jako ogół metod, zasad badań i procedur, stosowanych w tej lub innej dziedzinie naukowej;
- **metodologię stosowaną**, łączącą w sobie metodykę i technikę przeprowadzenia konkretnego badania.

Przeprowadzenie *badania naukowo-pedagogicznego* wymaga od badacza posiadania umiejętności posługiwania się wszystkimi poziomami metodologii całościowo, ponieważ one wzajemnie się dopełniają i konkretyzują. Przebywając w hierarchicznej zależności, każda kolejna metodologia (lub poziom metodologii) włącza w siebie wartości poprzedniej. Na przykład, metodologia konkretnej nauki zawiera zarówno problemy specyficzne dla naukowego poznania w danej dziedzinie (np. współzależność pedagogiki i psychologii), jak i problemy poprzednich poziomów (np. podejście systemowe i modelowanie w pedagogice, indukcyjne i dedukcyjne drogi poznania rzeczywistości pedagogicznej).

Podsumowując, można stwierdzić, że: **metodologię pedagogiki** należy zaliczyć do metodologii konkretno-naukowej i jest ona **nauką o ogóle metod, zasad badania i procedur, stosowanych w dziedzinie wiedzy pedagogicznej**.

Najczęściej metodologię pedagogiki traktuje się jako teorię metod badań pedagogicznych, a także jako teorię niezbędną dla powstania koncepcji edukacyjnych i wychowawczych. R. Barrou uważa, że istnieje filozofia pedagogiki, która opracowuje metodologię badania. W zakres jej zadań wchodzi również opracowywanie teorii pedagogicznej, logiki i sensu działalności pedagogicznej. Z tego punktu widzenia metodologia pedagogiki zawiera również filozofię edukacji, wychowania i rozwoju, jak i metody badania, pozwalające stwarzać teorie procesów i zjawisk pedagogicznych. Opierając się na tej przesłance, czeska pedagog-badaczka Jana Skalková stwierdza, że metodologia pedagogiki jest systemem wiedzy o podstawach i strukturze teorii pedagogicznej.

W celu pełnego ukazania istoty rozpatrywanego pojęcia należy podkreślić, że **metodologia pedagogiki**, oprócz już omówionych, wypełnia także inne funkcje:

- po pierwsze, określa sposoby otrzymywania wiedzy naukowej odzwierciedlającej stale zmieniającą się rzeczywistość pedagogiczną (M.A. Daniłow);

Podstawy metodologii pedagogiki

- po drugie, kieruje i warunkuje główny kierunek, za pomocą którego osiąga się konkretny cel naukowo-badawczy (P.W. Koppin);
- po trzecie, zabezpiecza wszechstronny odbiór informacji o badanym procesie lub zjawisku (M.N. Skatkin);
- po czwarte, pomaga wprowadzeniu nowej wiedzy do zasobów teorii pedagogiki (F.F. Korolow);
- po piąte, zabezpiecza uściślenie, wzbogacenie, systematyzację terminów i pojęć w nauce pedagogicznej (W.J. Gmurman);
- po szóste, stwarza system informacji, opierający się na faktach obiektywnych i logiczno-analitycznych narzędziach poznania naukowego (M.N. Skatkin).

Te cechy pojęcia „metodologia”, określające jej funkcje w nauce, pozwalają wyciągnąć wniosek, że **metodologia pedagogiki** jest konceptualnym przedstawieniem celu, treści, metod badania umożliwiających otrzymanie maksymalnie obiektywnej, dokładnej, usystematyzowanej wiedzy o procesach i zjawiskach pedagogicznych.

W związku z powyższym w **każdym badaniu pedagogicznym** można wyodrębnić **główne cechy metodologii**:

- po pierwsze, wyznaczenie celu badania z uwzględnieniem poziomu rozwoju nauki, potrzeb działalności praktycznej, aktualności społecznej i rzeczywistych możliwości naukowca lub grupy badaczy;
- po drugie, poznanie wszystkich procesów zachodzących w badaniu z punktu widzenia ich zależności wewnętrznych i zewnętrznych, rozwoju i samorozwoju.
- po trzecie, rozpatrzenie zagadnień edukacyjnych i wychowawczych z pozycji wszystkich nauk o człowieku: socjologii, psychologii, antropologii, fizjologii, genetyki. Wynika to stąd, że pedagogika jest nauką łączącą całą nowoczesną wiedzę ludzkości i wykorzystującą wszystkie dane naukowe o człowieku w interesach stworzenia optymalnych systemów pedagogicznych;
- po czwarte, ukierunkowanie na podejście systemowe w badaniu (struktura, wzajemny związek elementów i zjawisk, ich podporządkowanie, dynamika rozwoju, tendencje, istota i właściwości, czynniki i warunki);
- po piąte, ukazanie i rozwiązanie sprzeczności w procesie edukacji i wychowania, w rozwoju grupy lub osobowości;
- i, wreszcie, po szóste, opracowanie związków teorii i praktyki, idei i ich realizacji, ukierunkowanie pedagogów na nowe koncepcje naukowe, nowe myślenie pedagogiczne z jednoczesnym wyzbyciem się idei przestarzałych, przewycięciem istniejących w pedagogice postaw zachowawczych i konserwatywnych.

Obecnie szczególnie ważne znaczenie ma bynajmniej nie nowy problem podwyższenia jakości badań pedagogicznych. Metodologię ukierunkowuje się na okazanie pomocy pedagogowi-badaczowi przy formowaniu się jego specjalistycznych umiejętności w dziedzinie pracy badawczej. W ten sposób, metodologia otrzymuje normatywne ukierunkowanie, a jednym z jej ważniejszych zadań staje się metodologiczne zabezpieczenie pracy badawczej.

Metodologia pedagogiki jako dziedzina poznania naukowego występuje w dwóch aspektach: **jako system wiedzy i jako system działalności naukowo-badawczej**. Chodzi tu o dwa rodzaje działalności – badania metodologiczne i zabezpieczenie metodologiczne. Zadaniem badań metodologicznych jest ukazanie prawidłowości i tendencji rozwoju nauki pedagogicznej i jej związków z praktyką, zasad podniesienia jakości badań pedagogicznych, analiza zawartości ich systemu pojęć i wykorzystywanych metod. Natomiast zabezpieczenie badań metodologicznych oznacza wykorzystanie posiadanej wiedzy metodologicznej w celu uzasadnienia programu badania i oceny jego jakości w trakcie przeprowadzenia lub po zakończeniu takiego badania.

Ukazane różnice uwarunkowują wyodrębnienie dwóch funkcji metodologii pedagogiki – *deskrypcyjnej*, tj. opisowej, dopuszczającej także formowanie opisu teoretycznego obiektu, i *preskrypcyjnej*, tj. normatywnej, stwarzającej punkty orientacyjne dla pracy pedagoga-badacza. Obecność tych funkcji pozwala na wyodrębnienie dwóch grup – teoretycznych i normatywnych – podstaw metodologii pedagogiki.

Do **podstaw teoretycznych**, pełniących funkcje deskrypcyjne, zalicza się:

- definicję metodologii;
- ogólną charakterystykę metodologii nauki, jej poziomów;
- metodologię jako system wiedzy i system działalności, źródła zabezpieczenia metodologicznego działalności badawczej w dziedzinie pedagogiki;
- obiekt i przedmiot analizy metodologicznej w dziedzinie pedagogiki.

Podstawy normatywne rozpatrują następujący zakres problemów:

- poznanie naukowe w pedagogice wśród innych form duchowego poznania świata, do których należy odnieść poznanie wywołowo-empiryczne i artystyczno-obrazowe odzwierciedlenie rzeczywistości;
- charakter procesu wyznaczenia celów, wyodrębnienie specjalnego obiektu badań, zastosowanie specjalnych środków poznania, jednoznaczność pojęć;
- typologia badań pedagogicznych;
- charakterystyka badań, na podstawie których badacz może porównywać i oceniać własną pracę naukową w dziedzinie pedagogiki: problem, te-

mat, aktualność, obiekt badania, jego przedmiot, cel, zadania, hipoteza, bronione tezy, nowatorstwo, znaczenie dla nauki i praktyki;

- logika badania pedagogicznego itd.

Tymi podstawami określono obiektywny zakres badań metodologicznych. Ich wyniki mogą posłużyć jako źródło uzupełniania treści zarówno metodologii pedagogiki, jak i metodologicznej refleksji pedagoga-badacza.

W strukturze wiedzy metodologicznej E.G. Judin wyodrębnia cztery poziomy: **filozoficzny, ogólnonaukowy, konkretno-naukowy i technologiczny**.

Na zawartość pierwszego, wyższego poziomu filozoficznego metodologii, przede wszystkim składają się ogólne zasady poznania i kategorialny ustrój nauki. Funkcje metodologiczne pełni cały system wiedzy filozoficznej.

Drugi poziom, tj. metodologia ogólnonaukowa, zawiera koncepcje teoretyczne, stosowane we wszystkich lub w większości dziedzin naukowych.

Trzeci poziom – metodologia konkretno-naukowa, zawiera ogół metod, zasad badań i procedur, stosowanych w tej lub innej specjalistycznej dziedzinie naukowej. Metodologia konkretnej nauki zawiera zarówno problemy, specyficzne dla poznania naukowego w danej dyscyplinie, jak i kwestie pojawiające się na wyższych poziomach metodologii.

Na czwarty poziom – metodologia technologiczna – składa się metodyka i technika badań, tj. zestaw procedur niezbędnych dla otrzymania wiarygodnego materiału empirycznego i jego pierwotnej obróbki, po której materiał ten może być włączony do ogółu wiedzy naukowej. Na tym poziomie wiedza metodologiczna ma wyraźnie zaznaczony charakter normatywny.

Wszystkie poziomy metodologii pedagogiki tworzą z sobą system, w ramach którego między nimi istnieją określone współzależności. Warto w tym miejscu podkreślić, że poziom filozoficzny stanowi podstawową wartość założoną wszelkiej wiedzy metodologicznej, określając światopoglądowe podejście do procesu poznania i przekształcenia rzeczywistości.

Metodologia jako nauka o organizacji działalności opiera się na wiedzy naukowej. Pedagog-badacz, przystępując do działalności naukowej, powinien precyzyjnie uświadomić sobie, czym jest nauka, jej organizację, prawidłowości rozwoju, strukturę wiedzy naukowej. Pedagog-badacz powinien również wyraźnie rozumieć kryteria naukowości nowej wiedzy, którą planuje uzyskać, formy wiedzy naukowej, wykorzystywane przez niego w trakcie badań oraz takie, za pomocą których zamierza przedstawić wyniki badań naukowych, tj. wszystko to, na czym badacz będzie opierać się podczas pracy naukowo-badawczej, aby była ona uświadomiona i zorganizowana. Również włączając się w każdą innowacyjną pedagogiczną działalność edukacyjną, pedagog-praktyk będzie zmuszony zbadać wszystko, co dla realizacji

jego celów może zaoferować współczesna nauka. W związku z tym musi on również wyraźnie uświadomić sobie, jaka bywa wiedza naukowa, jak ona powstaje i jak jest wykorzystywana, jaką specyfikę ma pedagogika jako dziedzina wiedzy naukowej w aspekcie możliwości jej zastosowania w praktyce itd.

Struktura wiedzy naukowej. Strukturalizacja wiedzy naukowej przebiega na podstawie poszczególnych dziedzin nauki. W niniejszym opracowaniu autorzy ukazują specyfikę pedagogiki jako dziedziny wiedzy naukowej.

1. Pedagogika jest bezpośrednio związana ze wszystkimi naukami, ponieważ przygotowuje ona specjalistów dla wszystkich dziedzin nauki i dla wszystkich rodzajów działalności zawodowej.
2. Opiera się na wykorzystaniu osiągnięć prawie wszystkich nauk, przede wszystkim filozofii, psychologii, fizjologii, cybernetyki, socjologii.
3. Wyróżnia się nadzwyczajną złożonością obiektu badań, który, prawdopodobnie jest najbardziej skomplikowany w porównaniu z obiektami wszystkich innych nauk ze względu na fakt, że wpływa na niego całe mnóstwo najróżnorodniejszych czynników.
4. Podstawowe kategorie pedagogiki, takie jak „edukacja”, „wychowanie”, „rozwój” itd. wyjaśnia się bardzo różnie. Wiele pojęć pedagogiki pozostaje nieprzetłumaczalnych (ze względu na zawartość treściową) na pojęcia pedagogiki istniejące w innych językach – angielskim, niemieckim itd.
5. Pedagogika w pewnym sensie jest nauką o przyszłości, ponieważ obecnie proces edukacyjny powstaje opiera się na poglądach społeczeństwa na naukę i wychowanie we współczesnych warunkach; tylko w szkole uczniowie odbywają naukę przez 12 lat, następnie uczą się w jednej z wielu placówek kształcenia zawodowego od roku do pięciu lat, a potem mijają jeszcze około dziesięciu lat po ukończeniu edukacji, zanim były uczeń i/lub student uformuje się jako jednostka i profesjonalista. W ciągu całego tego okresu istotnie zmieniają się polityczne, społeczne i ekonomiczne warunki życia zarówno całego społeczeństwa, jak i pojedynczego człowieka.
6. Realizacja osiągnięć pedagogiki w praktyce edukacji i wychowania w znacznej mierze zależy od osobowości każdego pedagoga, jego przekonań, dążeń, profesjonalizmu.

1.1.1. Metody i środki badań

Środki i metody są bardzo ważnymi komponentami składowymi logicznej struktury organizacji działalności. Dlatego te składają się one na duży dział metodologii, jako nauki o organizacji działalności i odnoszą się nie

tylko do działalności naukowej, ale również do praktycznej działalności edukacyjnej itd. W trakcie rozwoju nauki powstają i doskonalą się środki poznania: materialne, matematyczne, logiczne, językowe. Oprócz tego, w ostatnim czasie pojawiły się całkowicie nowe, informatyczne środki poznania. Wszystkie te środki były stworzone specjalnie. W tym sensie materialne, informacyjne, matematyczne, logiczne, językowe środki poznania posiadają wspólną cechę: są konstruowane, tworzone, opracowywane dla tych lub innych celów poznawczych.

Materialne środki poznania – są to przede wszystkim narzędzia do badań naukowych, związane z empirycznymi metodami badań – obserwacją, pomiarem, eksperymentem. Środki te są bezpośrednio skierowane na badane obiekty, do nich należy główna rola w empirycznym sprawdzeniu hipotez oraz innych wyników badań naukowych, w odkryciu nowych obiektów, faktów.

Środki informatyczne. Masowe wprowadzenie do procesu edukacyjnego techniki i technologii informatycznej, środków telekomunikacji nie tylko w sposób zasadniczy przekształciło proces edukacyjny, ale również czyni z nich środki poznania pedagogicznego.

Matematyczne środki poznania. Rozwój matematycznych środków poznania, przenikających także do dziedzin humanistycznych, ma coraz większy wpływ na rozwój nowoczesnej nauki.

Pod wpływem matematycznych środków poznania zachodzą istotne zmiany w teoretycznym oprzyrządowaniu nauk opisowych. Środki matematyczne pozwalają usystematyzować dane empiryczne, ujawniać i formułować ilościowe zależności i prawidłowości. Środki matematyczne wykorzystuje się również jako specjalne formy idealizacji i analogii (modelowanie matematyczne). Obecnie w naukach opisowych, w tym także w pedagogice, największą rolę odgrywają środki statystyki matematycznej.

Środki logiczne. Podczas przeprowadzenia każdego badania naukowiec rozwiązuje zadania logiczne. Wykorzystanie środków logicznych w procesie budowy rozważań i zdobywania dowodów pozwala badaczowi oddzielać kontrolowane dowody od argumentów intuicyjnych lub przyjmowanych bezkrytycznie, błędne od prawdziwych, zamęt od sprzeczności.

Środki językowe. W tym językowym środku poznania są prawidłowości budowy pojęć (definicji). Podczas każdego badania naukowego uczyony zmuszony jest uściślać wprowadzone wcześniej pojęcia oraz używać nowych. Stwierdzenia zawsze są związane z językiem jako środkiem poznania i wyrażenia wiedzy.

W badaniach pedagogicznych istotną rolę odgrywa także wprowadzane przez uczoność korelacje języka pedagogiki ze specyficznymi cechami języków nauk pokrewnych – psychologii, socjologii.

1.2. Metodologiczne zasady badań

Rozpatrując to zagadnienie, opieramy się na założeniu, że metodologia jest niczym innym, jak zastosowaniem zasad ogólnych i teorii podczas rozwiązywania zadań badawczych, problemów konkretnej nauki. Jednocześnie należy zaznaczyć, że stopień wspólności samych zasad jest różny. Możemy mówić tylko o najbardziej ogólnych, tzw. powszechnych zasadach, prawach i kategoriach. Wszystkie one mają charakter filozoficzny i w danym wypadku dialektyka występuje jako ogólna metodologia poznania naukowego. Rozpatrzmy zatem podstawowe zasady metodologiczne badań psychologiczno-pedagogicznych.

Wielką rolę w pomyślnym urzeczywistnieniu badań psychologiczno-pedagogicznych odgrywa **zasada jedności teorii i praktyki**. Praktyka jest podstawowym kryterium prawdziwości tego lub innego założenia teoretycznego. Natomiast teoria, która nie opiera się na praktyce, jest kontemplacyjna, bezowocna. Zadanie teorii polega na ukazaniu drogi praktyce. Praktyka, nie poparta teorią naukową, jest skazana na wywołowość, bezcelowość, znikomą efektywność. Dlatego też przy organizowaniu badań psychologiczno-pedagogicznych ważne jest nie tylko opieranie się na osiągnięciach teorii psychologiczno-pedagogicznej, ale również uwzględnianie rozwoju praktyki. Bez głębokiej i wszechstronnej analizy naukowej działalności praktycznej przyszłych specjalistów nie sposób wytyczyć efektywnej drogi doskonalenia procesu edukacyjnego na wyuczonych uczelniach.

Jedną z zasad metodologicznych jest **twórcze, konkretno-historyczne podejście do badanego problemu**. Takiej postawy wymaga dialektyka. Doświadczenie wskazuje, że nie można dokładnie zbadać tego lub innego problemu przygotowania przyszłych specjalistów, wybierając wyłącznie szlaki wyznaczone wcześniej, działając szablonowo, nie wykazując inwencji twórczej.

W trakcie badań należy szukać własnych uargumentowanych wyjaśnień nowych faktów, zjawisk, dopełniać i uściślać ustalone przekonania, nie bać się wykazywać inwencji naukowej. Jednak inwencję tę należy połączyć z nauką zasadnością i przewidywalnością, ponieważ badania psychologiczno-pedagogiczne związane są z ludźmi, a każdy kontakt z człowiekiem powinien go wzbogacać duchowo. Twórczość jest nierozdzielnie związana z podejściem **konkretno-historycznym** do oceny zjawisk psychologiczno-pedagogicznych: to, co na pewnym etapie historycznym jest postępowe, w innych warunkach może być reakcyjne. Inaczej mówiąc, nie wolno oceniać teorii psychologiczno-pedagogicznych powstałych w przeszłości z perspektywy nowoczesności.

Twórcze podejście do rozwiązania badanego problemu jest ściśle związane z **zasadą obiektywności** rozpatrywania zjawisk psychologiczno-pedagogicznych. Sztuka badacza polega na tym, by znaleźć drogi i środki zrozumienia sedna zjawiska, jego świata wewnętrznego, nie wnosząc przy tym niczego zewnętrznego, subiektywnego.

Sukces badania psychologiczno-pedagogicznego w dużej mierze zależy od realizacji **zasady wszechstronności** badania procesów i zjawisk psychologiczno-pedagogicznych. Każdy fenomen pedagogiczny ma wiele powiązań z innymi zjawiskami i jego odizolowanie, rozpatrywanie jednostronne niechybnie doprowadzi do zniekształconych, błędnych wniosków.

Metodologiczna zasada wszechstronności zakłada **kompleksowe podejście do badania procesów i zjawisk pedagogicznych**. Jednym z najważniejszych wymagań podejścia kompleksowego jest ukazanie wszystkich wzajemnych powiązań badanego zjawiska, wszelkich mających na niego wpływ oddziaływań zewnętrznych, usunięcie wszystkich czynników przypadkowych, zniekształcających obraz badanego problemu. Innym istotnym wymaganiem takiego podejścia pozostaje wykorzystanie w trakcie badań różnorodnych metod oraz ich odmian.

Jedną z zasad metodologicznych badania psychologiczno-pedagogicznego jest jedność aspektów historycznego i logicznego. Logika poznania obiektu lub zjawiska ukazuje logikę jego rozwoju, tj. jego historię. Na przykład historia rozwoju osobowości jest swoistym kluczem do zrozumienia konkretnej osobowości, podjęcia praktycznych decyzji skierowanych na jej wychowanie i edukację. W historii rozwoju osobowości przejawia się jej istota, ponieważ człowiek tylko wtedy staje się osobowością, gdy posiada własną historię, drogę życiową, biografię.

Metodologiczną zasadą badania jest także **systemowość**, tj. systemowe podejście do badanych obiektów. Zasada ta zakłada rozpatrywanie obiektu badań jako systemu, ukazanie określonej złożoności jego elementów (ujawnienie i uwzględnienie wszystkich jest niemożliwe i niepotrzebne), określenie, klasyfikację i uporządkowanie zachodzących między tymi elementami związków, wyodrębnienie ze zbioru powiązań systemowych umożliwiających połączenie różnych elementów w jeden system.

Podejście systemowe ujawnia strukturę (ukazującą względną wywrotność) i organizację (ilościową charakterystykę i ukierunkowanie) systemu. Każdy element systemu w złożonych procesach może występować jako samodzielny system, a jego jakość zależy nie tylko od jakości poszczególnych elementów, ale także od relacji poszczególnych elementów ze środowiskiem.

Ważną rolę metodologiczną w badaniach psychologiczno-pedagogicznych odgrywają kategorie dialektyki: istota i zjawisko; przyczyna i następstwo; konieczność i przypadkowość; możliwość i rzeczywistość; treść i forma; jednostkowość, szczegół i ogół itd. Kategorie te w rękach pedagoga są niezastąpionym środkiem metodologicznym pozwalającym głęboko i wielostronnie rozpatrywać te problemy edukacji i wychowania przyszłych specjalistów.

W przypadku psychologii i pedagogiki fakt ten oznacza, że, po pierwsze, nawet unikatowe doświadczenie psychologiczno-pedagogiczne zawiera cechy, charakterystyczne dla każdego doświadczenia organizacji procesu edukacyjnego na uczelni wyższej, a po drugie, wszelkie ogólne tezy należy potwierdzać doświadczeniem, ukazywać w nich podłoże dla rozwoju i, wreszcie, po trzecie, nie ma i nie może być zaleceń, przydatnych na wszelkie wypadki życiowe.

Na podstawie zgromadzonych faktów odbywa się proces podniesienia poznania empirycznego do poziomu uogólnienia teoretycznego. Zachodzi przy tym charakterystyczny ruch od wiedzy jednostronnej do wielostronnej, na podstawie pierwotnych uogólnień powstają określone modele i idee, łączą się elementy zmysłowe i racjonalne, w trakcie których doznania zmysłowe i doświadczenia praktyczne uwalniają się od wszystkiego co przypadkowe i podnoszą się do poziomu teoretycznego, charakterystycznego dla szeregu podobnych zjawisk. Oczywiście, ważne jest, aby fakty były rozpatrywane w konkretnych uwarunkowaniach historycznych, całościowo, z uwzględnieniem ich wzajemnych powiązań. Tylko w takim przypadku będą one przekonywujące.

Konkretne etapy i sposoby zbioru, obróbki, uogólnienia i analizy materiału faktycznego wyjaśniają prawa logiki naukowej, będącej syntezą logiki dialektycznej i formalnej. Opanowanie myślenia naukowego pozostaje najważniejszą umiejętnością każdego badacza.

Należy podkreślić, że myślenie naukowe przede wszystkim zakłada swobodne operowanie przez badacza pojęciami i kategoriami naukowymi, zwłaszcza odnoszącymi się do tematu badania. Brak takich umiejętności uniemożliwia przeprowadzenie badań naukowych i swobodne poruszanie się w literaturze naukowej.

Ważne wymagania metodologiczne dotyczące badania problemów psychologiczno-pedagogicznych wynikają z podstawowych praw dialektyki, której sednem jest **prawo jedności i walki przeciwieństw**, przejawiające się w działaniu sprzeczności. Istnieją różne rodzaje sprzeczności: wewnętrzne i zewnętrzne, podstawowe i pochodne, główne i drugorzędne. Na przykład,

rozpatrując sprzeczności wewnętrzne i zewnętrzne w psychologii i pedagogice istnieje następująca klasyfikacja sprzeczności rozwoju jednostki.

Do pierwszej grupy sprzeczności, warunkujących rozwój osobowości przyszłego specjalisty, zazwyczaj zalicza się sprzeczności zachodzące między czynnikami zewnętrznymi. Szacunek do osoby jest podstawowym warunkiem pracy ka dego specjalisty. W związku z tym, spotykając się z jakimkolwiek przejawem ordynarności, brakiem nale ytej uwagi, biurokratyzmem ze strony kierownictwa, młodzi specjaliści nierzadko głęboko to prze ywają, co równie istotnie wpływa na rozwój ich osobowości. Jednak w większości przypadków sprzeczności zachodzące między czynnikami zewnętrznymi, warunkującymi rozwój jednostki, są siłą napędową jej harmonijnego formowania się i dojrzewania społecznego.

Do drugiej grupy sprzeczności zaliczane są zazwyczaj sprzeczności zachodzące między czynnikami zewnętrznymi i wewnętrznymi. Najwa niejsze spośród nich są: sprzeczności między wymaganiami w stosunku do jednostki a jej przygotowaniem do realizacji tych wymagań; między nowymi wymaganiami a ustalonymi przekonaniem i zwyczajowym zachowaniem; między zewnętrznymi wymaganiami a wymaganiami osoby w stosunku do samej siebie; między poziomem przygotowania osoby a mo liwościami wykorzystania posiadanej wiedzy, umiejętności i nawyków w praktyce.

Do trzeciej grupy sprzeczności zaliczane są sprzeczności zachodzące między czynnikami wewnętrznymi. U podstaw tych sprzeczności le y nierównomierność rozwoju poszczególnych przejawów, cech, komponentów strukturalnych osoby. Do tej grupy sprzeczności zalicza się tak e sprzeczności zachodzące między składnikiem racjonalnym świadomości a zmysłami; między oczekiwaniami jednostki a jej realnymi mo liwościami; między nowymi potrzebami a starymi stereotypami zachowania się; między nowym a starym doświadczeniem itd. Jednak podstawowym i głównym systemem sprzeczności wewnętrznych są sprzeczności zachodzące między motywami działalności, tworzącymi realną podstawę konkretnej jednostki i będącymi bezpośrednim przejawem systemu stosunków społecznych tworzących istotę tej jednostki. Badanie takiego systemu sprzeczności pozostaje wa nym zadaniem ka dego badania psychologiczno-pedagogicznego.

W rozwoju ka dego procesu i zjawiska zachodzą wielorakie sprzeczności. W związku z tym dokładnie zrozumienie jakiegokolwiek zjawiska, jego rozwoju jest niemo liwe bez konkretnej analizy systemu sprzeczności warunkujących ten rozwój.

Prawo przejścia zmian ilościowych w jakościowe wymaga badania wszystkich zjawisk psychologiczno-pedagogicznych w związku z ich cha-

rakterystyką jakościową i ilościową. Ka dy człowiek ma niezliczenie wiele ró norodnych cech, które mo na porównać z cechami innych ludzi. Jako całościowa określoność jakościowa człowiek jest istotą społeczną. Psychika człowieka posiada własną określoność jakościową. Jest ona pełna przykładów tzw. czystego, naocznego przejawu prawa przejścia zmian ilościowych w jakościowe. Na przykład, do pewnego poziomu minimalne bodźce receptorów (narządów odbierających bodźce zewnętrzne i wewnętrzne) nie doprowadzają do powstawania u człowieka odczuć, w ka dym razie na poziomie świadomym. Dopiero po przekroczeniu określonego poziomu (minimalny próg odczucia) dane bodźce zostają przyjęte świadomie, subiektywnie, odbywa się ich odzwierciedlenie na jakościowo innym poziomie.

Prawo zaprzeczenia zaprzeczenia jako prawo usuwania starego i wprowadzania nowego w procesie rozwoju postępowego, przy jednoczesnym zachowaniu niektórych elementów poprzedniego zjawiska lub procesu, ma wielorakie przejawy w yciu człowieka. Ka dy nowy etap w rozwoju osoby lub grupy w znaczeniu ściśle filozoficznym jest zaprzeczeniem starego, lecz zaprzeczenie to pozostaje częścią rozwoju postępowego. Wa ną rolę w takim zaprzeczeniu odgrywa samowychowanie jednostki, aktywna praca pedagoga ukierunkowana na formowanie osobowości przyszłego specjalisty.

Rola metodologiczna rozpatrzonych zasad, praw dialektyki przejawia się w konkretnym badaniu psychologiczno-pedagogicznym przede wszystkim poprzez logikę dialektyczną. W postaci skondensowanej wymagania logiki dialektycznej, wszystkich rozpatrzonych, a tak e innych zasad i kategorii dialektyki sprowadzają się do tego, by badać przedmiot wszechstronnie, śledząc jego rozwój, nie zapominając, e głównym kryterium weryfikującym zawsze pozostaje praktyka. Takie są najbardziej ogólne wymagania metodologiczne wobec konkretnego badania psychologiczno-pedagogicznego. Dialektyka, jej prawa, kategorie w konkretnym badaniu są uwzględniane, przede wszystkim, jako zasady powszechne.

Na podstawie zasad powszechnych ukształtowały się również bardziej szczegółowe wymagania zasadnicze, które badacz dziedziny psychologii i pedagogiki koniecznie powinien uwzględniać, np. zasada determinizmu; jedności oddziaływań zewnętrznych i wewnętrznych warunków rozwoju; aktywności osobowości; jedności psychiki i działalności; podejścia indywidualnego, społecznego, aktywnego i inne.

Zasada determinizmu nakłada na badacza obowiązek uwzględniania wpływu ró nych czynników na rozwój zjawisk psychologiczno-pedagogicznych. Podczas badania jednostki nale y uwzględniać niejako trzy podsystemy determinacji jej zachowania: przeszłość, terażniejszość i przyszłość, obiektywnie przez zachowanie odzwierciedlane.

Przeszłość jednostki pozostawia ślad zarówno na jej drodze życiowej, biografii, jak i cechach osobowych, moralnych i psychologicznych. Wpływ przeszłości, historii rozwoju jednostki na jej zachowanie nosi charakter pośredni. Bezpośredni zaś wpływ na zachowanie, postępowanie ma świadomość, motywy działalności jednostki. Działalność i kontakty z otoczeniem, a tak e warunki wewnętrzne rozwoju jednostki składają się na system jej determinacji do doskonalenia się.

Wyjątkowo wielki wpływ na rozwój jednostki mają cele jej działalności, które w znacznym stopniu ukierunkowane są na przyszłość. W tym sensie można mówić o przyszłości jako o podsystemie determinacji rozwoju jednostki. W takim wypadku uświadomiony cel jako prawo uwarunkowuje sposób, charakter działalności jednostki i przez to ma istotny wpływ na jej rozwój. Wszystkie trzy podsystemy (przeszłość, teraźniejszość i przyszłość) są determinowane, wzajemnie powiązane i wzajemnie uwarunkowane.

Zasada jedności oddziaływań warunków zewnętrznych i wewnętrznych. Zgodnie z tą zasadą poznanie życia wewnętrznego jednostki odbywa się w wyniku oceny zewnętrznych informacji o jej zachowaniu, działalności i czynach. W miarę rozwoju społecznego człowieka coraz bardziej zło ona staje się jego naturą wewnętrzną i cię ar właściwy warunków rozwoju, wewnętrznych w odniesieniu do zewnętrznych. Współzależność elementów wewnętrznych i zewnętrznych w rozwoju jednostki zmienia się zarówno historycznie, jak i na poszczególnych etapach drogi życiowej człowieka: im bardziej jest on rozwinięty, tym w większym stopniu postęp jego osobowości jest związany z aktualizacją czynników wewnętrznych.

Zasada aktywnej działalności jednostki akcentuje uwagę badacza na fakcie, że nie tylko otoczenie formuje jednostkę, ale tak e jednostka jest aktywnym obiektem poznania i przekształcenia otaczającego świata. Ta zasada zakłada rozpatrzenie wszystkich zachodzących w jednostce zmian przez pryzmat jej działalności. Wpływ działalności na jednostkę ma znaczenie ogromne. Poza działalnością człowiek nie istnieje, ale istota człowieka nie kończy się na działalności i nie może być do niej sprowadzana i z nią całkowicie utożsamiana. Oddziaływania psychologiczno-pedagogiczne na jednostkę muszą uwzględniać charakter jej działalności i nierzadko najbardziej efektywne oddziaływania polegają na zmianie, korekcie tej lub innej działalności człowieka.

Zasada rozwoju wskazuje, że rozpatrując zjawiska psychologiczno-pedagogiczne, należy uwzględniać ich zmienność, ruch i stałe dążenie do rozwiązywania sprzeczności pod wpływem systemu determinant wewnętrznych i zewnętrznych. Zasada rozwoju w psychologii i pedagogice zazwyczaj bywa

rozpatrywana w dwóch aspektach: rozwój historyczny jednostki od jej powstania do stanu współczesnego – filogeneza; i rozwój osobowości konkretnego człowieka – ontogeneza. Oprócz tego można i trzeba rozpatrywać rozwój różnych komponentów jednostki – ukierunkowanie, charakter, inne cechy osobowościowe. Oczywiście jest, że efektywność oddziaływań psychologiczno-pedagogicznych w decydującym stopniu zależy od tego, na ile kompletnie i dokładnie uwzględnia się rozwój przyszłego specjalisty, poddanego oddziaływaniu, na ile dokładnie uwzględnia się rozwój systemu pedagogicznego.

Konkretna realizacja wszystkich tych reguł odbywa się zgodnie z zasadą **podejścia indywidualnego, społecznego, aktywnego**. Ta zasada ukierunkowuje badacza na całościowe badanie jednostki z uwzględnieniem głównych czynników społecznych jej rozwoju – środowiska społecznego, wychowania, działalności osobowości, jej aktywności wewnętrznej.

Zasady pozostają bezpośrednią metodologią naukowych badań psychologiczno-pedagogicznych, uwarunkowują ich metodykę, wyjściowe koncepcje teoretyczne, hipotezy.

Opierając się na rozpatrzonych wyżej zasadach, spróbujmy sformułować wymagania metodologiczne dotyczące przeprowadzenia badań psychologiczno-pedagogicznych:

- a) procesy i zjawiska należy badać takimi, jakie one są, ze wszystkimi ich stronami pozytywnymi i negatywnymi, sukcesami i trudnościami, bez upiększania; prowadzić nie opis zjawisk, ale ich krytyczną analizę;
- b) operatywnie reagować na nowości w teorii i praktyce psychologii i pedagogiki;
- c) wzmacniać praktyczne ukierunkowanie, wagę i jakość zaleceń;
- d) mieć pewność prognozy naukowej, widzieć perspektywę rozwoju badanego procesu, zjawiska;
- e) kierować się ścisłą logiką myśli, dbać o przejrzystość eksperymentu psychologicznego lub pedagogicznego.

Na tej podstawie można określić wymagania metodologiczne odnośnie do wyników przeprowadzenia badań psychologiczno-pedagogicznych, które zostały przez nie uwarunkowane. Do nich zalicza się obiektywność, wiarygodność, pewność i słuszność.

1.3. Formy działalności poznawczej w dziedzinie pedagogiki

Nauka, mimo jej ogromnego znaczenia we współczesnym świecie nie jest jedyną formą jego poznania duchowego. Dla metodologii pedagogiki problem wyznaczenia specyfiki i miejsca poznania naukowego w odróżnieniu od innych form odzwierciedlenia rzeczywistości jest kwestią bardzo złożoną i przy tym wcale nie „czysto akademicką”. Spośród zarzutów wysuwanych pod adresem nauki pedagogicznej są: bezdusność, abstrakcyjność, pseudonaukowość. Zajmując stanowisko w tej sprawie, należy uwzględnić fakt, że działalność pedagogiczna pozostaje na tyle rozpowszechniona, wszechprzenikającą dziedziną bytu społecznego, że uzasadnione byłoby wyodrębnić ją w szczególną dziedzinę świadomości społecznej. Obecnie w dziedzinie tej można wskazać przynajmniej trzy formy odzwierciedlenia: 1) odzwierciedlenie rzeczywistości pedagogicznej w żywołowo-empirycznym procesie poznania; 2) artystyczno-obrazowe odzwierciedlenie rzeczywistości pedagogicznej; 3) odzwierciedlenie rzeczywistości pedagogicznej w poznaniu naukowym.

Opierając się na pracach W.I. Zagwiazńskiego, W.P. Kochanowskiego, W.W. Krajewskiego i innych badaczy, postaramy się ukazać zawartość wyżej wskazanych form odzwierciedlenia. Badacz opiera się na danych już zgromadzonych w danej nauce, zaczerpniętych z doświadczeń całej ludzkości. Najbardziej ogólnie naukę określa się jako dziedzinę działalności człowieka, której zadaniem jest opracowanie i systematyzacja teoretyczna obiektywnej wiedzy o rzeczywistości.

Działalność w dziedzinie nauki – badanie naukowe – pozostaje szczególną formą procesu poznania, systematycznym i ukierunkowanym badaniem obiektów za pomocą metod naukowych, zakończone formowaniem wiedzy o badanych obiektach.

Inną formą poznania jest poznanie żywołowo-empiryczne. W pedagogice często bywa tak, że te dwa rodzaje poznania – **naukowe i żywołowo-empiryczne** nie odróżniają się wystarczająco wyraźnie i uważa się, że pedagog-praktyk, nie stawiając przed sobą specjalnych celów naukowych i nie wykorzystując środków poznania naukowego, może nadal pełnić funkcję badacza. Często spotyka się pogląd, że wiedzę naukową można otrzymać w procesie praktycznej działalności pedagogicznej, bez poznania podstaw teoretycznych, tj. pokutuje pogląd, że teoria pedagogiczna nicomal sama „wyrasta” z praktyki. Jest to błędne założenie. Poznanie naukowe jest procesem

szczególnym. Składa się on z poznawczej działalności ludzi, środków poznania, jego obiektów i wiedzy. Przyjrzyjmy się zatem różnicom między poznaniem naukowym i ywiolowo-empirycznym.

Poznanie ywiolowo-empiryczne jest pierwotne. Istniało ono od zawsze i istnieje do dziś. W trakcie tego poznania odbiór wiedzy nie jest oddzielony od społeczno-praktycznej działalności człowieka. Źródłem wiedzy są różnorodne działania praktyczne z obiektami. Na podstawie własnego doświadczenia ludzie poznają cechy tych obiektów, poszukują najlepszych sposobów oddziaływania na nie – ich obróbkę, wykorzystanie.

W dziedzinie pedagogiki wiedza ywiolowo-empiryczna czerpana jest z pedagogiki ludowej. Mądrość ludowa przekazała nam wiele sprawdzonych rad pedagogicznych pod postacią przysłów i porzekadeł. W nich odzwierciedlają się określone prawidłowości pedagogiczne. Wiedzę takiego typu otrzymuje również nauczyciel w procesie pracy praktycznej z dziećmi.

Specyfika poznania naukowego, w odróżnieniu od poznania ywiolowo-empirycznego, przede wszystkim polega na tym, że działalność poznawczą w nauce prowadzą nie wszyscy, a jedynie specjalnie przygotowane grupy ludzi – pracownicy naukowcy. Formą jej urzeczywistniania i rozwoju staje się badanie naukowe.

W historii nauki powstały i opracowywane są nadal specjalne środki poznania, metody badania naukowego, podczas gdy poznanie ywiolowo-empiryczne takimi środkami nie dysponuje. Na przykład do zasobów środków poznania naukowego można zaliczyć: modelowanie, wykorzystywanie modeli idealizowanych, tworzenie teorii, hipotez, eksperymentowanie.

Nauka w odróżnieniu od ywiolowo-empirycznego procesu poznania bada nie tylko takie przedmioty, z którymi ludzie mają do czynienia podczas bezpośredniej praktyki, lecz również takie, które ujawniają się w trakcie rozwoju samej nauki. Nierzadko badanie tych przedmiotów poprzedza ich wykorzystanie praktyczne. W nauce zaczyna się specjalnie badać same rezultaty działalności poznawczej – wiedzę naukową. Opracowywane są kryteria, dzięki którym wiedzę naukową można oddzielić od wiadomości ywiolowo-empirycznych, od poglądów, od rozważań kontemplacyjnych, teoretycznych itd.

W odróżnieniu od takich nauk, jak matematyka, fizyka lub logika, nauka pedagogiczna korzysta z języka naturalnego, ze słownictwa powszechnie używanego. Jednak, kiedy słownictwo takie trafia do zasobów językowych nauki, powinno nabyć cechę charakterystyczną dla terminów naukowych, jaką jest jednoznaczność. Pozwala ona na jednolite rozumienie tych terminów przez wszystkich naukowców danej dziedziny. Gdy słowo z zasobów powszechnie używanego leksyki staje się terminem naukowym, zawiera ono

cię ar ogromnej pracy naukowej. Dlatego gdy słyszymy słuszne protesty przeciwko zaśmiecaniu wykładu naukowego, nie wolno ich traktować jako apelu do odrzucenia terminologii naukowej.

Upraszczenie terminologii naukowej często okazuje się niemożliwe dlatego, że termin to pewna formuła, za którą stoi wieloletnia praca naukowa, abstrahowanie, odkrycia. Próby zastąpienia „własnymi, prostymi, zrozumiałymi dla wszystkich słowami” terminu naukowego są skazane na porażkę, ponieważ doprowadziłyby to do stanu, gdy zamiast krótkiego zdania przyszłoby pisać całą księgę, a dokładniej, przepisywać, ponieważ opracowania, na podstawie których przyjęty został termin, już zostały napisane.

Proponowana przez badacza nowa wiedza naukowa wymaga odbioru aktywnego, a nie percepcji pasywnej. Bez samodzielnej interpretacji takiej wiedzy przez odbiorców pozostanie ona martwym, mało zrozumiałym zbiorem rozważań naukowych.

Istotnym brakiem w rozwoju nauki pedagogicznej jest niski poziom wymagań wobec jednoznaczności terminologicznej. Wiadomo, że stworzenie precyzyjnej i jednoznacznej terminologii jest jednym z najważniejszych wymagań stawianych metodologii naukowej. Problem ten nierzadko dotyczy podstawowych kategorii pedagogicznych, jak chociażby „wychowanie” i „edukacja”, które czasem bywają traktowane niejednoznacznie nawet w ramach jednej pracy naukowej. Empiryczna różnorodność w traktowaniu tych kategorii może stać się przeszkodą w badaniach nad teoretycznym i praktycznym połączeniem edukacji i wychowania. Wieloznaczność, dozwolona w języku potocznym, jest niedopuszczalna w nauce, zwłaszcza gdy dotyczy to poszczególnej koncepcji całościowej. W przypadku, gdy badacz stosuje (bez zastrzeżeń) niejednoznaczność terminologiczną w tej samej pracy, ma to bezpośredni wpływ na obniżenie jej jakości.

Nareszcie, kardynalna różnica między poznaniem naukowym a wywiadowo-empirycznym polega na tym, że badanie naukowe ma charakter systematyczny i ukierunkowane jest na rozwiązywanie problemów, które zostały świadomie sformułowane jako cel.

Jeżeli wiedza empiryczna została włączona w system nauki, traci swój charakter wywiadowy. Jeżeli obserwacje nad własną pracą lub nad pracą innych pedagogów-praktyków prowadzi w sposób ukierunkowany i systematyczny, wychodząc z pozycji naukowych, wykorzystując określone środki poznania naukowego, w takim wypadku otrzyma on materiał empiryczny, który można wykorzystać dla celów analizy teoretycznej. Jednak badacz, który dąży do tego, aby wszystkie konstrukcje teoretyczne wyprowadzać wyłącznie z doświadczeń, skazuje się na pracę mało wydajną, ponieważ poznanie empiryczne nie daje wiedzy o sednie.

Sformułowanie pojęcia o różnicy między powszechnym, ywiolowo-empirycznym, a naukowym procesem poznania jest dopiero pierwszym, chociaż bardzo ważnym punktem orientacyjnym dla oceny działalności w obszarze pedagogiki (z punktu widzenia przynależności procesu i rezultatów takiej działalności do dziedziny nauki). Podstawowe przejawy procesu poznania naukowego to: charakter procesu wyznaczenia celów, wyodrębnienie specjalnego obiektu badania, zastosowanie specjalnych środków poznania, jednoznaczność terminów. Jeżeli w jakiejś pracy, nawet bardzo interesującej i pożytecznej dla praktyki, brakuje wytycznych dotyczących sposobów pozyskania wiedzy o tym lub innym obiekcie pedagogicznym, a także brakuje pozostałych przejawów procesu poznania naukowego oznacza to, że takiej pracy nie można uważać za opracowanie naukowe.

Podsumowując, możemy stwierdzić, że **badania naukowe w dziedzinie pedagogiki** to specyficzny rodzaj działalności poznawczej, w trakcie której za pomocą różnorodnych metod ujawniają się nowe, wcześniej nieznanne strony i cechy badanego obiektu. Przy tym **główne zadanie badania** polega na ukazaniu związków wewnętrznych, prawidłowości i sił napędowych rozwoju procesów lub zjawisk pedagogicznych.

Każde badanie naukowe odbywa się na podstawie pewnych wskazówek metodologicznych. Metodologia charakteryzuje stosunek badacza do analizy rzeczywistości. Ona przenika w samo sedno badania, przejawia się w jego zamiarach, metodach i wynikach.

Według charakteru i zawartości, badania w dziedzinie pedagogiki dzielą się na fundamentalne, stosowane i opracowania.

Badania fundamentalne pozwalają rozwiązywać zadania o charakterze strategicznym. Ich cechami charakterystycznymi są: teoretyczna aktualność, przejawiająca się w ukazaniu prawidłowości, zasad lub faktów, mających zasadniczo ważne znaczenie, koncepcyjność, historyzm, analiza krytyczna naukowo bezpodstawnych założeń, wykorzystanie metodyk adekwatnych do natury poznawanych obiektów rzeczywistości, nowatorstwo i naukowa wiarygodność otrzymanych wyników. Jednak głównym kryterium fundamentalnego badania w dziedzinie pedagogiki jest rozwiązanie perspektywicznego zadania, jakim jest przygotowanie rozwoju nauki w ciągu najbliższych 10–15 i więcej lat, jak również takie rozważania teoretyczne, które wnoszą poważne zmiany w logikę rozwoju samej nauki.

Podstawowymi przejawami **badania stosowanych** są: ich bliskość do aktualnych potrzeb praktyki, porównawcze ograniczenie selekcji badania, operatywność w przeprowadzeniu i wprowadzeniu wyników itd. Rozwiązując operatywne zadania pedagogiki, badania stosowane opierają się na bada-

niach fundamentalnych, które dostarczają im narzędzi pomagających zorientować się w szczegółowych problemach, dają wiedzę teoretyczną i logiczną, pozwalają określić najbardziej racjonalną metodykę badania. Z kolei badania stosowane dają cenny materiał dla badań fundamentalnych.

Do **opracowań** w pedagogice zazwyczaj odnoszą się zalecenia metodyczne dotyczące wybranego problemu z zakresu edukacji i wychowania, instrukcje, pomoce i opracowania metodyczne. Opierają się one na badaniach stosowanych i nowatorskim doświadczeniu pedagogicznym. Charakterystycznymi cechami opracowań są: celowość, konkretność, jasność i stosunkowo niewielka objętość.

Specyficznym rodzajem badania naukowo-pedagogicznego jest dociekanie, uogólnienie i wprowadzenie w praktykę nowatorskich doświadczeń nauczania i wychowania. Cechą szczególną badań podobnego typu jest fakt, że zazwyczaj badania te są częścią konkretnej praktyki pedagogicznej i dostępne są dla każdego nauczyciela.

1.3.1. Nauka jako instytucja społeczna

System placówek naukowych Ukrainy składa się z setek instytutów akademii nauk, instytutów badawczych i ośrodków akademii edukacji, akademii nauk medycznych, akademii rolniczych i tysięcy naukowo-badawczych instytutów bran oowych.

Znaczna część potencjału naukowego w każdym kraju zawsze była skupiona na wyższych uczelniach. Fakt ten można z jednej strony wyjaśnić tym, że w celu zabezpieczenia wysokiego poziomu nauczania w szkołach wyższych niezbędna jest wysoko wykwalifikowana kadra naukowo-pedagogiczna. Z drugiej strony, pozwala to wciągać młodych naukowców prosto z ław studenckich do badań naukowych. W wyższych placówkach oświatowych – uniwersytetach, akademiach i instytutach pracuje, w zależności od liczebności studentów od kilkuset do kilku tysięcy osób kadry profesorsko-naukowej. Oprócz tego, praca naukowa odbywa się również w instytutach doskonalenia zawodowego, na przykład w znajdujących się w różnych regionach Ukrainy instytutach i akademiach doskonalenia pracowników oświaty. Żadna praca naukowa nie może odbywać się bez odpowiedniej infrastruktury. Są to tzw. instytucje wspomagania naukowego: wydawnictwa naukowe, czasopisma naukowe, zakłady budowy oprzyrządowania naukowego itd., które występują w roli niejako podbranych nauki jako instytucji społecznej.

Nauka jako instytucja społeczna może funkcjonować wyłącznie pod warunkiem posiadania specjalnie przygotowanej wykwalifikowanej kadry naukowej. Przygotowanie naukowej (naukowo-pedagogicznej) kadry na Ukrainie odbywa się poprzez studia doktoranckie, w wyniku których kandydaci otrzymują **stopień naukowy doktora nauk** [kandydat nauk].

Osoby, które otrzymały tytuł doktora nauk, po ukończeniu przewodu habilitacyjnego, stanowią naukową (naukowo-pedagogiczną) kadre o wyszej kwalifikacji – **doktorzy habilitowani** [doktor nauk].

Jednocześnie ze stopniami naukowymi wykładowcy w wyszych placówkach oświatowych, otrzymują **tytuły naukowe**, które odzwierciedlają stopień ich kwalifikacji pedagogicznych: **docenta** (zazwyczaj są to doktorzy nauk posiadający stałą pracę wykładowcy na uczelni wyszej i opublikowane prace naukowe) i **profesora** (zazwyczaj są to doktorzy habilitowani posiadający duże prace naukowe – podręczniki, monografie).

Obecnie wiele szkół średnich zatrudnia pracowników naukowo-pedagogicznych z uczelni wyszych lub instytucji naukowych. Fakt, że w szkołach ogólnokształcących, liceach, w placówkach kształcenia zawodowego pracuje coraz więcej doktorów i doktorów habilitowanych, świadczy o tym, że również te placówki oświatowe coraz częściej będą prowadzić działalność naukowo-badawczą.

1.3.2. Cechy indywidualnej działalności naukowej

1. Pracownik naukowy powinien wyraźnie rozgraniczać ramy swojej działalności i określić cele własnej pracy naukowej. W nauce, tak samo jak w każdej innej dziedzinie działalności zawodowej, odbywa się naturalne rozgraniczenie pracy. Pracownik naukowy nie może zajmować się „nauką w ogóle”, nie może nawet zajmować się pedagogiką „w ogóle”, a powinien wyodrębnić wyraźny i wąski kierunek pracy, postawić konkretny cel i konsekwentnie dążyć do jego osiągnięcia.

Jedną z głównych cech pracownika naukowego jest zdolność skupienia się tylko na tym problemie, którym się zajmuje, a wszystkie pozostałe, „poboczne” kwestie, musi wykorzystywać tylko w tej mierze i na tym poziomie, na jakim są one opisane w istniejącej na dzień dzisiejszy literaturze naukowej.

2. Praca naukowa opiera się na osiągnięciach poprzedników. Przed przystąpieniem do pracy naukowej mającej na celu rozwiązanie jakiegoś pro-

blemu, należy zbadać całą literaturę przedmiotu, zapoznać się z osiągnięciami w danej dziedzinie dokonanymi przez poprzedników.

3. Pracownik naukowy powinien opanować terminologię naukową i ściśle przestrzegać jej u ywania. W ka dej nauce mo na zaobserwować zjawisko równoległego istnienia ró nych szkół naukowych. Na przykład, edukacją problemową zajmują się szkoły M.I. Machmutowa, I.J. Lerniera, T.W. Kudriawcewa (mówiąc o szkole naukowej, nale y mieć na uwadze nieformalną grupę, stworzoną przez wybitnych uczonych). Ka da szkoła naukowa ma własny aparat pojęciowy.
4. Wyniki ka dej pracy naukowej, ka dego badania powinny obowiązkowo przybrać formę pisemną – naukowe sprawozdanie, odczyt, referat, artykuł, książ ka itd.

Oprócz tego, liczba i objętość publikacji naukowych są równie wskaźnikiem, wprawdzie formalnym, produktywności ka dego pracownika naukowego i ka dy badacz stale prowadzi i uzupełnia spis swoich opublikowanych prac.

I.4. Metoda poznania naukowego

Działalność człowieka w ka dej jej formie (naukowa, praktyczna itd.) wyznacza cały szereg czynników. Ostateczny jej wynik zale y nie tylko od tego, kto działa (podmiot) lub na co ona jest ukierunkowana (obiekt), ale równie od tego, jak przebiega dany proces, jakie sposoby, środki były wykorzystane. Właśnie to składa się na problemy metody.

Metoda (grec. – sposób poznania) – „droga do czegokolwiek”, sposób osiągnięcia celu, w określony sposób uporządkowana działalność podmiotu w ka dej jej formie.

Podstawowa funkcja metody – wewnętrzna organizacja i uregulowanie procesu poznania lub praktycznego przekształcenia jakiegoś obiektu. Otó pojęcie metody (w ka dej jej formie) sprowadza się do ogółu określonych przepisów, sposobów, norm poznania i działania. Jest ona systemem zaleceń, zasad, wymagań, które powinny ukierunkowywać badacza na rozwiązanie konkretnego problemu, osiągnięcie określonego rezultatu w ka dej sferze działalności. Metoda dyscyplinuje poszukiwanie prawdy, pozwala (je eli jest poprawna) zaoszczędzić wysiłek i czas, poruszać się do celu najkrótszą drogą. Prawdziwa metoda jest rodzajem swoistego kompasu, dzięki któremu podmiot poznania i działania wyznacza własną drogę, pozwalającą na uniknięcie błędów.

Pojęcie „**metoda naukowa**” należy rozumieć jako „ukierunkowane podejście, drogę, za pośrednictwem której osiąga się postawiony cel. Jest to zestaw różnych poznawczych sposobów i operacji praktycznych, ukierunkowanych na nabycie wiadomości naukowych” (Kochanowski 1999). W psychologii i pedagogice metoda naukowa rozumiana jest jako system sposobów, odpowiednich dla zawartości i zadań danej nauki.

Pojęcie „metoda” stosuje się w szerokim i wąskim znaczeniu tego słowa. W szerokim znaczeniu oznacza ono proces poznawczy, w którym zawarto kilka sposobów. Na przykład metoda analizy teoretycznej zawiera także syntezę, abstrahowanie, uogólnienie itd. W wąskim znaczeniu „metoda” oznacza specjalne sposoby dyscypliny naukowej. Na przykład w psychologii i pedagogice są to: metoda obserwacji naukowej, metoda ankietowania, metoda eksperymentalna i inne.

Na każdym etapie rozwoju nauki znaczenie metody poznania było wysoko oceniane przez wszystkich badaczy. Na przykład Francis Bacon porównywał metodę z pochodnią, oświetlającą podróżnikowi drogę w ciemności, i był zdania, że nie można liczyć na sukces w rozwiązywaniu jakiegokolwiek problemu, idąc fałszywą drogą. Filozof ten dążył do stworzenia takiej metody, która byłaby „organem” (narzędziem) poznania, zapewniającym człowiekowi przewagę nad przyrodą. W charakterze takiej metody widział on indukcję, która wymaga od nauki opierania się na empirycznej analizie, obserwacji i eksperymentach, aby na tej podstawie można było poznać przyczyny i prawa.

Kartezjusz nazywał metodą „dokładne i proste zasady”, których przestrzeganie sprzyja wzrostowi wiedzy, pozwala odróżnić zafałszowane od prawdziwego. Uczony ten mówił, że lepiej nie rozmyślać o odnajdywaniu różnych prawd, niż czynić to bez wykorzystania jakiegokolwiek metody, a zwłaszcza, bez dedukcyjno-racjonalistycznej.

Istotny wkład w metodologię poznania naukowego wniosła niemiecka filozofia klasyczna (Georg Hegel) i materialistyczna (Karol Marks), którzy dosyć głęboko opracowali metodę dialektyczną – odpowiednio na podstawie idealistycznej i materialistycznej.

W.P. Kochanowski uważa, że „każda metoda okazuje się nieefektywną, a nawet bezużyteczną, jeżeli będzie ona wykorzystywana nie jako «nieprzewodnia» w naukowej lub innej formie działalności, a jako gotowy szablon do «przemalowywania» faktów. Głównym przeznaczeniem każdej metody jest, w oparciu o odpowiednie zasady (wymagania, zalecenia itp.), zapewnić pomyślnie rozwiązanie określonych problemów poznawczych i praktycznych, wzrost wiedzy, optymalne funkcjonowanie i rozwój tych lub innych obiektów”.

Biorąc pod uwagę powyższe rozważania, możemy wyciągnąć następujące wnioski:

1. Metoda, z zasady, stosowana jest nie w izolacji, a w połączeniu, współdziałaniu z innymi metodami. Oznacza to, że ostateczny wynik działalności naukowej w dużej mierze zależy od tego, w jakim stopniu umiejętnie i efektywnie wykorzystano „w sprawie” heurystyczny potencjał każdej ze stron tej lub innej metody i wszystkie je razem. Każdy element metody istnieje nie sam przez się, lecz jako część całości i stosuje się jako całość. Szczególne znaczenie ma zdolność opanowania przeciwnych podejść metodologicznych i ich poprawne połączenie.
2. Powszechną podstawą systemu wiedzy metodologicznej jest filozofia jako metoda uniwersalna. Jej zasady, prawa i kategorie określają ogólny kierunek i strategię badania, „wypełniają” wszystkie inne poziomy metodologii, w sposób szczególny przenikają się wzajemnie i materializują się w konkretnej formie na każdym z tych poziomów. W badaniu naukowym nie można ograniczać się tylko zasadami filozoficznymi, nie wolno również zostawiać ich poza uwagę, jako coś nieodpowiadającego naturze danej działalności. Oczywiście, jeżeli przez pojęcie filozofia będziemy rozumieć poszukiwanie wiedzy w jej najogólniejszej, najszerzej formie, to filozofię można uważać za „matkę wszystkich poszukiwań naukowych”. Historia poznania i praktyki potwierdziły ten wniosek.
3. W swoim zastosowaniu każda metoda modyfikuje się w zależności od konkretnych warunków, celów badania, charakteru rozwiązywanych problemów, cech obiektu, tej lub innej dziedziny zastosowania metody (przyroda, społeczeństwo, poznanie), specyfiki badanych prawidłowości, osobliwości zjawisk i procesów (materialne lub duchowe, obiektywne lub subiektywne). W związku z powyższą treścią systemu metod, wykorzystywanych do rozwiązania określonego problemu, zawsze jest konkretna, ponieważ w każdym wypadku treść jednej metody lub systemu metod modyfikuje się zgodnie z naturą badanego procesu.

1.4.1. Klasyfikacja metod poznania naukowego

Różnorodność rodzajów działalności człowieka warunkuje różnorodne spektrum metod, które mogą być klasyfikowane według najróżniejszych założeń (kryteriów). Przede wszystkim należy wyodrębnić metody działalności duchowej, idealnej (w tym także naukowej) i metody działalności materialnej (praktycznej). Obecnie stało się oczywiste, że system metod, metodologia, nie

mogą być ograniczone tylko do dziedziny poznania naukowego, lecz powinny przekraczać jej zasięg i koniecznie włączać w zakres swojego zainteresowania również dziedzinę praktyki. Równocześnie należy uwzględnić ściśle współdziałanie tych dwóch dziedzin działalności człowieka.

Co zaś się tyczy metod nauki, to podstaw do ich podziału na grupy mogą być kilka. Otóż, w zależności od roli i miejsca w procesie poznania naukowego można wyodrębnić metody **formalne i treściowe, empiryczne i teoretyczne, fundamentalne i stosowane, metody badania i wykładu** itp. Zawartość badanych przez naukę obiektów służy jako kryterium do różnicowania metod przyrodoznawstwa i metod nauk społeczno-humanistycznych. Z kolei metody nauk przyrodniczych mogą być podzielone na metody badania przyrody nieożywionej i metody badania przyrody ożywionej. Wyodrębnia się również metody jakościowe i ilościowe, jednoznaczno-deterministyczne i probabilistyczne, metody poznania bezpośredniego i pośredniego, oryginalne i pochodne itd.

Do charakterystycznych cech metody naukowej (obojętnie do jakiego typu ona się odnosi) najczęściej zalicza się: obiektywność, odtworzenie, heurystyka, konieczność, konkretność i inne.

W nowoczesnej nauce wykorzystuje się wielopoziomową koncepcję wiedzy metodologicznej. W tym przypadku wszystkie metody poznania naukowego, zdaniem W.P. Kochanowskiego, można podzielić na następujące podstawowe grupy (według stopnia wspólnoty i zakresu ich zastosowania).

1. Metody filozoficzne, wśród których najstarszymi są metoda dialektyczna i metafizyczna. W gruncie rzeczy każda koncepcja filozoficzna ma funkcję metodologiczną, jest szczególnym sposobem działalności umysłowej. Dlatego te metody filozoficzne nie ograniczają się do dwóch wymienionych. Do nich zalicza się także metody, jak analityczna, intuicyjna, fenomenologiczna, hermeneutyczna.

Metody filozoficzne wskazują wyłącznie najogólniejsze kierunki badania, jego strategię generalną, ale nie zastępują metod specjalnych i nie określają ostatecznego rezultatu poznania wprost i bezpośrednio. Doświadczenie pokazuje, że im bardziej ogólna jest metoda poznania naukowego, tym bardziej jest ona nieokreślona odnośnie do zalecenia konkretnych kroków poznania, tym większa jest jej niejednoznaczność w stwierdzeniu ostatecznych wyników badania.

2. Ogólnonaukowe sposoby i metody badania niejako występują w charakterze swoistej „metodologii pośredniej” między filozofią a fundamentalnymi teoretyczno-metodologicznymi założeniami nauk specjalnych. Do zakresu pojęć ogólnonaukowych najczęściej zalicza się takie pojęcia, jak „informa-

cja”, „model”, „struktura”, „funkcja”, „system”, „element”, „optymalność”, „prawdopodobieństwo” i inne.

Cechami charakterystycznymi pojęć ogólnonaukowych są, po pierwsze, „przenikanie” w ich treści poszczególnych cech, przejawów, pojęć pojedynczych nauk i kategorii filozoficznych. Po drugie, mo liwość formalizacji, uściślenia za pomocą środków teorii matematycznej, logiki symbolicznej.

Je eli kategorie filozoficzne zawierają w sobie maksymalnie mo liwy stopień wspólnoty, tj. czynnik konkretno-powszechny, to dla pojęć ogólnonaukowych bardziej właściwy jest czynnik abstrakcyjno-ogólny (jednakowy). Pozwala to ukazać pojęcia ogólnonaukowe za pomocą środków abstrakcyjno-formalnych.

Na podstawie pojęć i koncepcji ogólnonaukowych formułowane są odpowiednie metody i zasady poznania, które zabezpieczają związek i optymalne współdziałanie filozofii ze specjalistyczną wiedzą naukową i jej metodami. Do zasad i sposobów ogólnonaukowych zalicza się: systemowo-jednostkowy i strukturalno-funkcjonalny, cybernetyczny, probabilistyczny, modelowanie, formalizację i szereg innych.

Wa na rola pojęć ogólnonaukowych polega również na tym, e dzięki swojemu „charakterowi pośredniemu” zapewniają one pośrednie przenikanie się wiedzy filozoficznej i wiedzy poszczególnych nauk (jak również odpowiednich metod). Chodzi o to, e wiedza filozoficzna nie nakłada się w czysto zewnętrzny, bezpośredni sposób na wiedzę poszczególnych nauk. W związku z tym próby przeniesienia wprost specjalnej naukowej treści na język kategorii filozoficznych są, z zasady, niekonstruktywne i mało efektywne.

3. Metody poszczególnych nauk – jest to ogół metod, zasad poznania, sposobów i procedur badawczych, stosowanych w tej lub innej nauce. Są to metody mechaniki, fizyki, chemii, biologii i nauk społeczno-humanistycznych. Omówione w dalszej części niniejszego opracowania metody badań psychologiczno-pedagogicznych odnoszą się do metod poszczególnych nauk.

4. Metody dyscyplinarne – jest to system metod, stosowanych w tej lub innej dyscyplinie naukowej, wchodzącej w jakąkolwiek dziedzinę nauki lub powstałej na pograniczu nauk. Ka da fundamentalna nauka jest kompleksem dyscyplin mających własny specyficzny przedmiot i własne metody badań.

5. Metody badań międzydyscyplinarnych – jest to ogół szeregu syntetycznych, integracyjnych metod (powstałych w wyniku połączenia elementów różnych poziomów metodologii), ukierunkowanych w głównej mierze na pogranicza dyscyplin naukowych. Szerokie zastosowanie metody te znalazły w realizacji kompleksowych badań i programów naukowych.

Metodologia zatem nie może być sprowadzona do jakiejś jednej, nawet bardzo ważnej metody. Naukowiec nigdy nie powinien polegać na jakiejś jednej nauce, nigdy nie powinien ograniczać metody swojego myślenia do jednej filozofii. Zdaniem W.P. Kochanowskiego „Metodologia nie jest tak prostą sumą poszczególnych metod, ich «mechaniczną jednością» lecz złożonym, dynamicznym, całościowym, subordynowanym systemem sposobów, zasad różnych poziomów, zakresów działania, ukierunkowania, możliwości heurystycznych, zawartości, struktur itd.”.

1.4.2. Klasyfikacja metod badań psychologiczno-pedagogicznych

Obecnie zostały określone niektóre podejścia do budowy systemu metod psychologii i pedagogiki. Najogólniej wszystkie one zazwyczaj dzielą się według stopnia wspólności na trzy grupy:

1. **Powszechna metoda badania naukowego** – dialektyczna. W tym punkcie system metod psychologii i pedagogiki jest identyczny z metodologią i metodami kadego poznania naukowego. Metoda dialektyczna jako podstawa wszystkich innych metod badań immanentnie jest w nich obecna.
2. Metody badań stosowane w psychologii i pedagogice, ogólne dla szeregu nauk. Metody te często nazywane są **ogólnonaukowymi logicznymi metodami poznania**.
3. **Szczegółowe metody psychologii i pedagogiki**, właściwe wyłącznie dla tych nauk.

Podobna klasyfikacja metod znajduje uznanie zarówno w psychologii, jak i w pedagogice. O to, stopień wspólności (ogólne – szczególne – pojedyncze) występuje jako wspólne kryterium dla klasyfikacji metod naukowych. Jednak w związku z tym, że metoda jest sposobem osiągnięcia określonych wyników w poznaniu i praktyce i że zawsze zawiera ona dwie organicznie związane części – obiektywną i subiektywną, w psychologii i pedagogice szczególne metody badań przyjęto klasyfikować według całego szeregu założeń.

Na przykład E.I. Monoszon (*Введение в научное исследование по педагогике*, 1988), opierając się na założeniu, że każde badanie psychologiczno-pedagogiczne realizowane jest na trzech poziomach – empirycznym, teoretycznym i metodologicznym, proponuje odpowiednio klasyfikować metody badania.

Badacz ten jest zdania, że na pierwszym poziomie kształtują się nowe fakty nauki i na podstawie ich uogólnienia formułowane są prawidłowości empiryczne. Na drugim poziomie pojawiają się i formułują podstawowe,

ogólne prawidłowości pedagogiczne, pozwalające wyjaśnić wcześniej odkryte fakty, a także przepowiedzieć i przewidzieć przyszłe. Na trzecim, metodologicznym, poziomie na podstawie badań empirycznych i teoretycznych formułowane są ogólne zasady i metody badania zjawisk pedagogicznych, budowania teorii. Wychodząc z powyższego, autor wyodrębnia empiryczne, teoretyczne i metodologiczne metody badań psychologiczno-pedagogicznych.

W.I. Zagwiazinski uważa, że metody badań psychologiczno-pedagogicznych można zgrupować według różnych cech (Zagwiazinski, Atachanow 2001). Na przykład, biorąc pod uwagę cechy przeznaczenia, w jednym przypadku, można wyodrębnić metody zbioru materiału faktycznego, jego interpretacji teoretycznej, skierowanego przekształcenia. W innym zaś przypadku, wyodrębnia się metody diagnostyki, objaśnienia, prognozowania, korekcji, obróbki statystycznej materiału i inne. Jednocześnie, biorąc pod uwagę poziom przeniknięcia w istotę badanych zjawisk psychologiczno-pedagogicznych, uczyony wyodrębnia dwie grupy metod – badania empirycznego i teoretycznego. Pierwsza grupa metod opiera się na doświadczeniu, praktyce, eksperymencie itp., a druga jest związana z abstrahowaniem od rzeczywistości zmysłowej, budowaniu modeli itd.

Podobne podejście do klasyfikacji metod badań psychologiczno-pedagogicznych można spotkać w pracach W.P. Dawydowa (Dawydow 1997). Między innymi uważa on, że w psychologii i pedagogice dzielenie metod badań na empiryczne i teoretyczne jest bardzo względne. Chodzi o to, że w systemie metodycznym tych nauk takie ogólnonaukowe logiczne metody badań (E.I. Monosz nazywa je metodologicznymi metodami badań), jak analiza i synteza, indukcja i dedukcja, porównanie, klasyfikacja itp., organicznie splatają się z metodą konwersacji, eksperymentu, analizy wyników działalności oraz innymi tradycyjnymi metodami tych nauk. W wielu podręcznikach psychologii i pedagogiki ogólnonaukowe logiczne metody badań wcześniej nawet nie były rozpatrywane. Dopiero w ostatnich monografiach, podręcznikach i pomocach naukowych zaczęto akcentować konieczność rozsądnego stosowania danych metod w badaniach psychologiczno-pedagogicznych. Wynika to z faktu, że logiczna kultura badania znacznie podnosi jego skuteczność. Czynnikiem ten jest zwłaszcza widoczny w czasie, kiedy w nowoczesnej psychologii i pedagogice stosowane są złożone metody badania, jest wykorzystywana analiza wieloaspektowa, metoda analogii, metoda formalizacji, metoda modelowania i inne. W związku z tym, że rozpatrywane metody logiczne, sposoby myślenia są niejako częścią składową tradycyjnych metod psychologiczno-pedagogicznych, próba odniesienia wielu z tych tradycyjnych metod do grupy empirycznych jest dość względna. W praktyce

metody te przewyższają konstatacje empiryczne, proste ukazania tych lub innych faktów psychologiczno-pedagogicznych, ponieważ zawierają w sobie element analizy teoretycznej.

Opierając się na wnioskach W.P. Dawydowa, warunkowo można wyodrębnić grupy empirycznych i teoretycznych metod badań psychologiczno-pedagogicznych. Umowność ta pozwala głębiej zorientować się w sednie systemu metod, podnieść kulturę pracy badawczej.

Do grupy teoretycznych metod badań W.P. Dawydow zalicza: teoretyczną analizę i syntezę, abstrahowanie i konkretyzację, indukcję i dedukcję, metodę modelowania. Natomiast w grupie empirycznych metod badań znalazły się: obserwacja, konwersacja, metody ankietowe (ankietowanie, wywiad, testowanie, socjometria), eksperyment i inne.

Oprócz wymienionych grup metod, W.P. Dawydow i W.I. Zagwiazinski uważają za mało wyodrębnienie w osobną grupę pomocniczych metod badań psychologiczno-pedagogicznych, do których zaliczają matematyczne i statystyczne metody interpretacji rezultatów pracy naukowej.

Jednocześnie, klasyfikacja zaproponowana przez W.P. Dawydowa wyróżnia się wyodrębnieniem w osobną grupę metod porównawczo-historycznych badań psychologiczno-pedagogicznych: genetycznych, historycznych i porównawczych. Badacz słusznie uważa, że zjawiska i procesy pedagogiczne można na wszechstronnie zrozumieć wyłącznie pod warunkiem dokładnego zbadania ich stanu współczesnego i poprzedniego rozwoju, ogólnych i szczegółowych cech w konkretno-historycznych warunkach.

Dalszą analizę metod badań psychologiczno-pedagogicznych będziemy prowadzić z uwzględnieniem właśnie takiej klasyfikacji.

1.4.3. Ogólnonaukowe logiczne metody i sposoby poznania

W związku z tym, że ogólnonaukowe logiczne metody poznania *a priori* wykorzystywane są w każdym badaniu psychologiczno-pedagogicznym, niezależnie od tego czy jest ono przeprowadzane na poziomie teoretycznym czy empirycznym, należy w tym miejscu przedstawić ich istotę bardziej szczegółowo. Przypomnijmy, że do nich zalicza się metody analizy, syntezy, abstrahowania, idealizacji, uogólnienia, indukcji, dedukcji i analogii.

Podkreślimy od razu, że w niniejszym opracowaniu przedstawione metody poznania będą rozpatrzone z pozycji ogólnometodologicznej. Postaramy się również ukazać istotę empirycznego i teoretycznego poziomu badania

w psychologii i pedagogice, koncentrując uwagę na szczegółach zastosowania wymienionych metod w konkretnych warunkach.

Opierając się na podejściu zaproponowanym przez W.P. Kochanowskiego (Kochanowski 1999), postaramy się krótko scharakteryzować każdą spośród ogólnonaukowych, logicznych metod poznania.

1. Analiza (grec. – rozłożenie, rozbiór) – rozłożenie obiektu na komponenty w celu ich samodzielnego zbadania. W encyklopediach pedagogicznych można spotkać następującą definicję analizy: *analysis* – badanie każdego elementu lub cechy zjawiska jako części całości, rozłożenie badanego przedmiotu lub zjawiska na składniki, wyodrębnienie w nim poszczególnych cech. Analizę można scharakteryzować jako proces rozłożenia, rozdzielania przedmiotów i zjawisk na niektóre cechy (części) w celu ich zbadania. Jednak takie podejście nie zakłada wyjaśnienia, ukazania i badania tej podstawy całości, która łączy wszystkie cechy, części przedmiotu, zjawiska w całość. Zadanie analizy polega na tym, by z różnego rodzaju danych, niekiedy dość luźnych, objaśniających niektóre zjawiska i fakty, przedstawić ogólny całościowy obraz procesu, ujawnić właściwe mu prawidłowości, tendencje.

Zwrócenie się do literatury filozoficznej daje nam jeszcze kilka pojęciowych znaczeń, w których z różnych stron rozpatrywana jest istota pojęcia „analiza” (B.A. Biriukow, I.G. Gierasim, M.K. Mordaszwili i inni):

- używa się jej w logice i teorii poznania w znaczeniu podstawowego logicznego środka, podstawowej operacji logicznej, uzyskania wiedzy poprzez wyłącznie logiczną obróbkę danych;
- jako metoda badania w ogóle, stanowiąca podstawę naukowej dialektycznej metody poznania;
- jako metoda badania i opisu formalnych zasad i struktur, wykorzystywanych przez naukę;
- jako sposób uzyskania wiedzy o przedmiocie w trakcie jego badania, jako specjalny środek badania zjawisk.

Na szczególną uwagę zasługuje charakterystyka analizy z pozycji dialektyki, gdzie jest ona traktowana jako specjalny środek badania zjawisk i uzyskiwania teoretycznej wiedzy o tych zjawiskach. Podstawowe poznawcze zadanie analizy dialektycznej polega na tym, by z różnorodności cech badanego przedmiotu wyodrębnić jego istotę (nie poprzez mechaniczne rozłożenie całości na części, ale poprzez wyodrębnienie i zbadanie cech podstawowej sprzeczności w przedmiocie), ukazać podstawę, łączącą wszystkie jego cechy w całość i wyprowadzić na tej podstawie prawidłowość rozwijającej się całości.

W pedagogice analiza występuje jako metoda lub sposób poznania rzeczywistości pedagogicznej.

Analizę stosuje się zarówno w działalności rzeczywistej (praktyka), jak i w myślowej. Rozróżnia się następujące rodzaje analizy: mechaniczny rozbiór; stwierdzenie składu dynamicznego; ukazanie form współdziałania elementów całości; znajdowanie przyczyn zjawisk; wykazanie poziomów wiedzy i jej struktury itp. Analiza nie powinna pomijać jakości przedmiotów. W każdej dziedzinie wiedzy znajduje się swoista granica rozłożenia obiektu, po której przekroczeniu przechodzimy do innego świata właściwości i prawidłowości (atom, molekula itp.). Odmianą analizy jest także rozbiór kategorii (zbiorów) przedmiotów na podkategorie – klasyfikacja i periodyzacja.

2. Synteza (grec. – połączenie, łączenie, układanie) to połączenie realne lub myślowe różnych cech, części przedmiotu w jedną całość. W ten sposób, syntezę należy rozumieć jako proces praktycznego lub myślowego jednoczenia całości z części lub połączenie różnych elementów, cech przedmiotu w jedną całość, niezbędny etap poznania. Jednocześnie należy wziąć pod uwagę, że synteza nie jest dowolnym, eklektycznym połączeniem „wyrwanych” części, „kawałeczków” całości, ale dialektyczną całością z wyodrębnieniem istoty. Współczesną naukę charakteryzuje nie tylko wewnątrz-, ale i międzydyscyplinarna synteza, a także synteza nauki oraz innych form świadomości społecznej. Wynikiem syntezy jest całkowicie nowy twór, którego cechą jest nie tylko zewnętrzne połączenie właściwości komponentów, ale także wynik ich wewnętrznych związków wzajemnych i współzależności.

Analiza i synteza są dialektycznie powiązane. Odgrywają one ważną rolę w procesie poznawczym i zachodzą na wszystkich jego poziomach.

3. Abstrahowanie jako metoda naukowego poznania. „Abstrakcja (łac. – oddzielenie, odłączenie, oderwanie) – a) aspekt, moment, część całości, fragment rzeczywistości, coś nierozwiniętego, jednostronnego, fragmentarycznego (abstrakcyjnego); b) proces myślowego wyodrębniania z szeregu cech i związków badanego przedmiotu lub zjawiska z jednoczesnym wyodrębnieniem właściwości interesujących w danym momencie podmiot (abstrahowanie); c) wynik abstrahującej działalności myślenia (*abstrakcja w wąskim znaczeniu*)”. Za pomocą abstrakcji powstały wszystkie pojęcia logiczne. Są to różnego rodzaju „przedmioty abstrakcyjne”, którymi mogą być zarówno osobno wzięte pojęcia i kategorie („rozwój”, „myślenie” itp.), jak i ich systemy (najbardziej rozwiniętymi spośród nich są matematyka, logika i filozofia).

Wyjaśnienie, które z rozpatrywanych właściwości są istotne, a które drugorzędne, pozostaje głównym problemem abstrahowania. Odpowiedź na pytanie, co w obiektywnej rzeczywistości wyodrębnia abstrahująca czyn-

ność myślenia, od czego myślenie odrywa się, w każdym konkretnym przypadku zależy przede wszystkim od natury badanego przedmiotu lub zjawiska, a także od zadań poznania. W trakcie swojego rozwoju historyczna nauka przechodzi od jednego poziomu abstrakcyjności do drugiego, wyższego.

Istnieją różne rodzaje abstrakcji:

- **abstrakcja utożsamienia**, w wyniku której wyodrębnia się ogólne cechy i zależności badanych metod. Tworzone są odpowiadające tym metodom kategorie powstałe na ustanowieniu równości przedmiotów z danymi cechami lub zależnościami, uwzględnia się czynnik identyczności w przedmiotach i odbywa się abstrahowanie od wszystkich różnic między nimi;
- **abstrakcja izolująca** – akty tzw. czystej abstrakcyjności, pozwalającej na wyodrębnienie pewnych właściwości i stosunków, które zaczyna się traktować jako samodzielne indywidualne przedmioty („przedmioty abstrakcyjne” – „dobroć”, „empatia” itp.);
- **abstrakcja aktualnej nieskończoności w matematyce** – powstaje, kiedy zbiory nieskończone rozpatruje się jak skończone. W tym wypadku badacz abstrahuje od zasadniczej niemożliwości utrwalenia i opisanego elementu nieskończonego zbioru, przyjmując takie zadanie jako skończone;
- **abstrakcja potencjalnej wykonalności** – opiera się na założeniu, że można urzeczywistnić każdą, ale skończoną liczbę operacji w procesie działalności matematycznej.

Abstrakcje różnią się także według poziomów (kategorii). Abstrakcje wychodzące od przedmiotów realnych nazywane są abstrakcjami pierwszego poziomu. Abstrakcje wychodzące od abstrakcji pierwszego poziomu nazywane są abstrakcjami drugiego poziomu itd. Najwyższym poziomem abstrakcji charakteryzują się kategorie filozoficzne.

4. Idealizacja jako metoda poznania naukowego najczęściej rozpatrywana jest jako specyficzny rodzaj abstrahowania. Idealizacją jest myślowe konstruowanie pojęć o obiektach nieistniejących i niemożliwych do wykonania w rzeczywistości, ale takich, dla których istnieją pierwowzory w świecie realnym.

W procesie idealizacji zachodzi maksymalne oddalenie od wszystkich realnych właściwości przedmiotu z jednoczesnym wprowadzeniem do zawartości tworzonych pojęć cech nieistniejących w rzeczywistości. W rezultacie powstaje tzw. obiekt idealizowany, którym można operować myślenie teoretyczne przy odzwierciedleniu realnych obiektów.

W wyniku idealizacji powstaje taki model teoretyczny, w którym charakterystyki i cechy poznawanego obiektu (przedmiotu, zjawiska) nie tylko są

oddalone od faktycznego materiału empirycznego, ale równie poprzez myślowe wyodrębnienie występują w postaci bardziej gwałtownej i pełnej, niżeli w rzeczywistości. Ostatecznie idealizowany obiekt występuje jako odzwierciedlenie realnych przedmiotów i procesów. Utworzywszy za pomocą idealizacji w takiego typu obiektach konstrukcje teoretyczne, można na równie później operować nimi w rozważaniach tak jak z przedmiotem realnie istniejącym i budować schematy abstrakcyjne rzeczywistych procesów, niezbędne do głębszego ich zrozumienia.

W ten sposób idealizowane przedmioty nie występują jako czyste fikcje, niemające niczego wspólnego z rzeczywistością, ale są wynikiem dość złożonego, pośredniego jej odbicia. Idealizowany obiekt ukazuje w poznaniu rzeczywiste przedmioty, jednak nie według wszystkich, a tylko dokładnie wyznaczonych cech. Obiekt taki jest uproszczonym i schematyzowanym obrazem rzeczywistego przedmiotu. Z zasady teoretyczne twierdzenia bezpośrednio odnoszą się nie do rzeczywistych obiektów, ale do obiektów idealizowanych, z którymi poznawcza działalność pozwala ustalać zasadnicze związki i prawidłowości, niedostępne podczas badania rzeczywistych obiektów, rozpatrywanych w całej ich różnorodności cech i związków empirycznych. Idealizowane obiekty są wynikiem różnorodnych eksperymentów myślowych, skierowanych na realizację pewnego niewykonalnego w rzeczywistości przypadku.

5. Uogólnienie, po pierwsze, jako metoda poznania naukowego występuje w formie logicznego procesu przejścia od jedności do ogółu, od mniej ogólnej do bardziej ogólnej wiedzy, ustalenia ogólnych cech i przejawów przedmiotów; po drugie, jako rezultat tego procesu: uogólnione pojęcie, opinia, prawo, teoria. Otrzymanie uogólnionej wiedzy pozwala głębiej pokazać rzeczywistość, zrozumieć jej istotę. Rozróżnia się dwa rodzaje uogólnień naukowych: wyodrębnienie wszystkich cech (abstrakcyjno-ogólne) lub tylko istotnych (konkretno-ogólne, tj. prawo).

Według innych założeń można wyodrębnić uogólnienia: a) od poszczególnych faktów, zdarzeń do ich wyrażenia w myślach (uogólnienie indukcyjne); b) od jednej myśli do innej, bardziej ogólnej myśli (uogólnienie logiczne). Przejście myślowe od bardziej ogólnego do mniej ogólnego składa się na proces ograniczenia. Uogólnienie nie może być bezgraniczne. Jego granicą są kategorie filozoficzne, które nie mają pojęcia rodzajowego i dlatego nie można na ich uogólnić.

6. Indukcja (łac. – wprowadzenie) to logiczna metoda (sposób) badania, związana z uogólnieniem wyników obserwacji i eksperymentów, a także procesem myślowym skierowanym od jedności do ogółu. W indukcji wyniki

doświadczenia „naprowadzają” na ogół, indukują go. Ponieważ doświadczenie zawsze pozostaje nieskończone i niepełne, wnioski indukcyjne zawsze mają charakter problematyczny (probabilistyczny). Indukcyjne uogólnienia zazwyczaj rozpatrywane są jako prawdy doświadczone albo prawa empiryczne. W słowniku języka polskiego pojęcie indukcja tłumaczone jest jako sposób rozważań polegający na „wyprowadzaniu wniosków ogólnych z przesłanek będących szczególnymi przypadkami tych wniosków”.

W.P. Kochanowski wyodrębnia następujące rodzaje uogólnień indukcyjnych:

- A. **Indukcja popularna** występuje wtedy, gdy regularnie powtarzające się cechy, obserwowane u pewnych przedstawicieli badanego zbioru (klasy) i utrwalane w przesłankach wniosku indukcyjnego, przenoszone są na wszystkich przedstawicieli badanego zbioru (klasy), w tym również na niezbadane jego części. Otóż to, co jest słuszne w „n” obserwowanych przypadków, jest słuszne również w następnym lub we wszystkich podobnych obserwowanych przypadkach. Jednak wyciągnięty wniosek często okazuje się fałszywy (na przykład „wszystkie białe łabędzie”), co następuje wskutek pośpiesznego uogólnienia. W ten sposób, omawiany rodzaj uogólnienia indukcyjnego występuje dopóty, dopóki nie natrafi na przypadek, przeczący mu (na przykład fakt obecności czarnych łabędzi). Popularną indukcję nierzadko nazywa się również indukcją przez wyliczenie (indukcja **enumeracyjna**).
- B. **Indukcja niezupełna** występuje wtedy, gdy pojawia się stwierdzenie, że wszystkim przedstawicielom badanego zbioru (klasy) przysługuje właściwość „n” na podstawie tego, że „n” charakteryzuje pewnych przedstawicieli tego zbioru. Na przykład fakt, że niektóre metale mają właściwość przewodzenia prądu elektrycznego, oznacza, że wszystkie metale przewodzą prąd elektryczny.
- C. **Indukcja zupełna** występuje wtedy, gdy pojawia się stwierdzenie, że wszystkim przedstawicielom badanego zbioru przysługuje cecha „n” na podstawie otrzymanej podczas badania doświadczonej informacji o tym, że każdy przedstawiciel badanego zbioru posiada właściwość „n”.

Rozpatrując indukcję zupełną, należy pamiętać, że, po pierwsze, nie daje ona nowej wiedzy i nie przekracza granic tego, co znajduje się w jej przesłankach. Niemniej jednak, ogólne wnioski otrzymane na podstawie badania pojedynczych przypadków, sumują zawartą w nich informację, pozwalają ją uogólnić, usystematyzować. Po drugie, chociaż wnioski z indukcji zupełnej mają w większości przypadków wiarygodny charakter, jednak nie są one wolne od błędów. Błędy te powstają głównie w następstwie

opuszczenia jakiegoś pojedynczego przypadku (czasem świadomie, umyślnie, aby udowodnić swoją słuszność), wskutek czego wnioski nie dotyczą wszystkich przypadków i przez to są nieuzasadnione.

- D. **Indukcja naukowa** występuje wtedy, gdy oprócz formalnego uzasadnienia uogólnienia otrzymanego drogą indukcyjną, przedstawia się treściowe dodatkowe uzasadnienie jego prawdziwości, w tym również za pomocą dedukcji (teorii, praw). Indukcja naukowa jest źródłem wiarygodnych wniosków dzięki temu, że tu szczególną uwagę zwraca się na związki konieczne, słuszne i przyczynowe.

W każdym badaniu naukowym należy ustalić związki przyczynowo-skutkowe zachodzące między różnymi przedmiotami i zjawiskami. W tym celu wykorzystywane są odpowiednie, bazujące na wnioskach indukcyjnych metody. Spróbujmy w tym miejscu rozpatrzyć podstawowe **metody indukcyjne ustanowienia związków przyczynowych** (zasady badania indukcyjnego Bacona–Milla).

- a) **Metoda jedynej zgodności** mówi: jeżeli obserwowane przypadki jakiegoś zjawiska mają wspólną tylko jedną przyczynę, to prawdopodobnie jest ona przyczyną danego zjawiska.

Inaczej mówiąc, jeżeli poprzedzające okoliczności „ABS” powodują zjawiska „abs”, a przyczyny „ADE” zjawiska „ade”, to można wyciągnąć wniosek, że „A” jest przyczyną „a” (lub zjawisko „A” i „a” są związane przyczynowo).

Zastosowanie metody jedynej zgodności w rzeczywistym badaniu naukowym napotyka poważne problemy, po pierwsze, w związku z tym, że w wielu przypadkach niełatwo oddzielić różne zjawiska jedno od drugiego. Po drugie, ogólną przyczynę należy wstępnie odgadnąć lub założyć, zanim przystąpi się do odszukiwania jej wśród różnych czynników. Po trzecie, bardzo często przyczyna nie sprowadza się do jednego ogólnego czynnika, a zależy od innych przyczyn i warunków. Dlatego aby zastosować metodę jedynej zgodności, należy mieć już określoną hipotezę o możliwej przyczynie zjawiska i zbadać wiele różnych zjawisk, w których zachodzi pewne działanie (skutek), w celu zwiększenia stopnia potwierdzenia wysuwanej hipotezy itd.

- b) **Metoda jedynej różnicy** mówi: jeżeli okoliczności, przy których zjawisko zachodzi lub nie zachodzi, różni się tylko jedną poprzednią okolicznością, a wszystkie inne okoliczności są identyczne, to ta jedna okoliczność jest przyczyną danego zjawiska.

Inaczej mówiąc, jeżeli poprzedzające okoliczności „ABS” powodują zjawiska „abs”, a okoliczności „BC” (zjawisko „A” usuwane jest w trak-

cie eksperymentu) powodują zjawiska „bc”, to mo na wyciągnąć wniosek, e „A” jest przyczyną „a”. Podstawą takiego wniosku jest fakt zniknięcia „a” przy usuwaniu „A”.

- c) **Połączona metoda zgodności i różnicy** powstaje jako potwierdzenie wyniku, otrzymanego za pomocą metody jedynej zgodności, poprzez zastosowanie wobec niego metody jedynej różnicy, tj. jest to kombinacja pierwszych dwu metod.
- d) **Metoda zmian współtowarzyszących** mówi: je eli zmiana jednej okoliczności zawsze powoduje zmianę innej, to pierwsza okoliczność jest przyczyną drugiej. Przy tym pozostałe poprzedzające zjawiska pozostają niezmiennie. Inaczej mówiąc, je eli przy zmianie poprzedzającego zjawiska „A” zmienia się również obserwowane zjawisko „a”, a pozostałe poprzedzające zjawiska pozostają niezmiennie, to mo na wyciągnąć wniosek, e „A” jest przyczyną „a”.

Rozpatrzone metody ustanowienia związków przyczynowych najczęściej stosowane są nie pojedynczo, ale dopełniają jedna drugą. Przy tym nie wolno popełniać błędów: „po tym z powodu tego”.

7. Dedukcja (łac. – wyprowadzenie): – po pierwsze, jest to przejście w procesie poznania od ogółu do jedności (szczegółu), wyprowadzenie szczegółu z ogółu; po drugie, jest to proces rozumowania logicznego, tj. przejścia według tych lub innych zasad logiki od niektórych danych założeń (przesłanek) do wniosków. Jako jedna z metod (sposobów) poznania naukowego dedukcja jest ściśle związana z indukcją. Są to swoiście dialektycznie związane sposoby ruchu myśli. W.P. Kochanowski uważa, że wielkie odkrycia, skoki myśli naukowej do przodu tworzone były właśnie przez indukcję – metodę chociaż ryzykowną, to rzeczywiście twórczą. Oczywiście nie należy sądzić, że rygorystyczność rozumowania dedukcyjnego nie ma żadnej wartości. W rzeczywistości właśnie dedukcja chroni wyobraźnię przed popełnieniem błędów, tylko ona pozwala, po ustanowieniu przez indukcję nowych punktów wyjściowych, wyciągnąć wnioski i zestawiać konkluzje z faktami. Dedukcja daje również narzędzia do sprawdzenia hipotez i jest cenną odtrutką na ponad miarę wybujałą fantazję.

8. Analogia (grec. – zgodność, podobieństwo) jest metodą poznania naukowego, dzięki której mo na ustalić podobieństwo niektórych cech, stanów i związków między nieidentycznymi obiektami. Wnioskowanie przez analogię – to wnioski wyciągnięte na podstawie takiego podobieństwa. W ten sposób, przy wnioskowaniu przez analogię wiedza otrzymana po rozpatrzeniu jakiegoś obiektu („modelu”), będzie przenosić się na inny, mniej zbadany i mniej dostępny dla badacza obiekt. Wnioski wyciągnięte przez

analogię są prawdopodobnymi: na przykład, jeżeli na podstawie podobieństwa dwu obiektów według pewnych parametrów wyciąga się wniosek o ich podobieństwie również według innych parametrów. Analogia nie daje wiedzy wiarygodnej: jeżeli przesłanki rozważania przez analogię są prawdziwe, nie oznacza to jeszcze, że również jego wnioski będą prawdziwe. Dla podniesienia prawdopodobieństwa wniosków wyciągniętych przez analogię należy dążyć do osiągnięcia następujących warunków:

- a) należy wychwycić wewnętrzne, a nie zewnętrzne właściwości zestawianych obiektów;
- b) obiekty te muszą mieć podobne cechy ważne i istotne, a nie przypadkowe i drugorzędne;
- c) liczba podobnych cech powinna być jak największa;
- d) należy uwzględniać nie tylko podobieństwa, ale także różnice, aby tych ostatnich nie przenieść na inny obiekt.

9. Modelowanie jako metoda poznania naukowego jest oparta na odtworzeniu cech pewnego obiektu na innym obiekcie, specjalnie stworzonym dla ich zbadania. Ten ostatni obiekt nazywany jest modelem przez co należy rozumieć obiekt posiadający pewne podobieństwa z prototypem i wykorzystywany jako środek opisu i/lub objaśnienia, i/lub prognozowania zachowania się prototypu (Kukuszkin 1998). Potrzeba zastosowania modelowania zachodzi wtedy, gdy badanie samego obiektu jest niemożliwe, trudne, drogie, czasochłonne itd.

Między modelem a oryginałem powinno istnieć znane podobieństwo (stosunek podobieństwa): cech, funkcji fizycznych; zachowania się badanego obiektu i jego opisu matematycznego; struktury itd. Właśnie to podobieństwo pozwala przenosić informację, otrzymaną w wyniku badania modelu, na oryginał.

Formy modelowania są różnorodne i zależą od wykorzystywanych modeli i dziedziny zastosowania modelowania. Według cech modeli wyodrębnia się modelowanie materialne (przedmiotowe) i idealne, przedstawione w odpowiedniej formie znakowej. Modele materialne są obiektami naturalnymi, podlegającymi w swoim funkcjonowaniu naturalnym prawom fizyki, mechaniki itp. Przy modelowaniu fizycznym (przedmiotowym) konkretnego obiektu jego badanie zastępowane jest badaniem pewnego modelu, mającego taką samą fizyczną naturę, jak oryginał (modele samolotów, statków itp.). Podczas modelowania idealnego (znakowego) modele mają postać schematów, wykresów, rysunków technicznych, wzorów, systemów równań. Do modelowania idealnego zalicza się, tzw. modelowanie myślowe, które jest klasyfikowane na modelowanie poglądowe, symboliczne i matematyczne.

Modelowanie poglądowe odbywa się na podstawie wyobrażeń badacza o realnym obiekcie za pomocą stworzenia modelu poglądowego, odzwierciedlającego zjawiska i procesy zachodzące w obiekcie. Z kolei modelowanie poglądowe można podzielić na hipotetyczne, analogowe i makietowane.

Podczas **modelowania hipotetycznego** zakłada się hipotezę o prawidłowościach przebiegu procesów w obiekcie realnym, która odzwierciedla poziom wiedzy badacza o obiekcie i opiera się na związkach przyczynowo-skutkowych między wejściem i wyjściem badanego obiektu.

Modelowanie analogowe opiera się na zastosowaniu analogii różnego poziomu. Z zasady model analogowy odzwierciedla kilka lub tylko jedną stronę funkcjonowania obiektu.

Modelowanie makietowane jest związane ze stworzeniem makiety realnego obiektu w określonej skali i jego zbadanie.

Modelowanie symboliczne – jest sztucznym procesem stworzenia logicznego obiektu, który zastępuje realny i odzwierciedla jego podstawowe cechy za pomocą określonego systemu znaków i symboli. Modelowanie symboliczne w zależności od stosowanych jednostek semantycznych przyjęto dzielić na językowe (opisowe) i znakowe (graficzne).

Modelowanie matematyczne bazuje na opisie realnego obiektu za pomocą narzędzi matematycznych. Obecnie najbardziej rozpowszechnione jest (komputerowe) modelowanie obiektów.

Podczas rozpatrywania powyższego zagadnienia zostały omówione tylko najistotniejsze metody ogólnonaukowego poznania logicznego, które znalazły szerokie zastosowanie zarówno na empirycznym, jak i na teoretycznym poziomie przeprowadzenia badań psychologiczno-pedagogicznych.

Rozdział II

Metody badań pedagogicznych

Do rozwiązania problemów badawczych w pedagogice wykorzystywane są metody badań naukowo-pedagogicznych. Im bogatszy zasób metod, tym większe sukcesy działalności naukowców. W związku z tym zasoby instrumentarium naukowego pedagogiki ciągle są uzupełniane zarówno poprzez konstruowanie nowych metod, jak i poprzez zapożyczenia przydatnych dla pedagogiki metod w innych naukach.

Specyficzną cechą nowoczesnej pedagogiki jest tendencja przekształcania metod poszukiwania naukowego w metody działalności praktycznej. Przyczyna tego zjawiska związana jest z odnową modeli dydaktycznych, z wprowadzeniem do praktyki metodyk edukacji badawczej, z koniecznością sprawowania kontroli nad poziomem wiedzy i wychowania dzieci.

Metodą badania naukowo-pedagogicznego nazywa się sposób otrzymywania wiarygodnej informacji o zjawisku pedagogicznym; sposób budowania teorii pedagogicznej.

Podstawowymi wymaganiami stawianymi przed doбором i zastosowaniem metod badań naukowo-pedagogicznych są: *kompleksowość* (dla uzyskania wiarygodnych danych należy korzystać z kompleksu metod, niejednokrotnie potwierdzających otrzymaną informację) i *adekwatność* wybranych metod do istoty badanego zjawiska i otrzymywanego rezultatu działalności pedagogicznej.

Mimo dużej liczby metod wykorzystanych w badaniach naukowo-pedagogicznych problem ich klasyfikacji praktycznie pozostaje nierozwiązany. Najczęściej w literaturze pedagogicznej można spotkać wyliczeniowy opis znanych metod, rzadziej – podział na określone grupy. W niniejszym opracowaniu jako podstawę analizy metod badań naukowo-pedagogicznych przyjęto ich podział na trzy grupy: *empiryczne*, *teoretyczne* i *statystyczno-matematyczne* (Babanski 1982).

2.1. Badania empiryczne w pedagogice

U podstaw empirycznych metod badań naukowo-pedagogicznych leży *doświadczenie praktyczne* (empiryka). Celem tych metod jest zbiór i pierwotna obróbka danych o zjawisku pedagogicznym, a także wykorzystanie uzyskanych wyników w praktycznej działalności pedagoga. Metoda ta zakłada bezpośredni kontakt badacza z badanym.

Do **empirycznych metod** badań naukowo-pedagogicznych zalicza się:

- obserwację;
- metody wywiadu: ustnego (rozmowę, wywiad), pisemnego (ankietowanie, testowanie);
- metodę uogólnienia niezależnych opinii, „konsylium” pedagogiczne;
- badanie dokumentacji szkolnej i wyników twórczej działalności uczniów;
- metodę badania monograficznego zjawiska pedagogicznego.

1. Obserwacja – najbardziej dostępna i rozpowszechniona metoda badania praktyki pedagogicznej. U jej podstaw leży *sposstrzegawczość*, jako zawodowa cecha pedagoga, umiejętność nie tylko patrzenia, ale także widzenia tego, co mu potrzebne. Cecha ta opiera się na *uwadze*, która bywa *rozproszona* (umiejętność widzenia jednocześnie całego szeregu obserwowanych zjawisk, nie koncentrując się na szczegółach [jak praca fotografa]) i *koncentrowana* (umiejętność wyławiania szczegółów, drobiazgów, złożoności badanego zjawiska [jak praca zegarmistrza lub jubilera]). Pedagog powinien jednocześnie być i „fotografem”, i „jubilerem”, tj. perfekcyjnie opanować obydwie formy uwagi. Na przykład, obserwując całą klasę, powinien on widzieć również każdego ucznia z osobna. Człowiek nie posiada takich cech od urodzenia, w związku z czym zachodzi konieczność stałego i ukierunkowanego ich trenowania.

Obserwacja pedagogiczna może być **powszednia**, z której nauczyciel korzysta w codziennej działalności praktycznej, i **naukowa**, niezbędna dla przeprowadzenia pracy naukowo-badawczej.

Obserwacją naukową jest specjalnie zorganizowane postrzeganie zjawiska lub procesu pedagogicznego, badanego w warunkach naturalnych.

Obserwacja naukowa istotnie odróżnia się od powszedniej. Przeprowadzając ją, należy wcześniej określić zadania, wyznaczyć obiekty, wykonać schemat obserwacji, określić warunki jej przeprowadzenia (długotrwałość, miejsce, czas, wykonawców itd.). Wyniki obserwacji naukowej obowiązkowo należy utrwalić w sprawozdaniach lub notatkach w dzienniku, zestawie

w tablicach, a jeżeli istnieje taka możliwość, można wykorzystać nagrania audio i wideo. Informacje, otrzymane w wyniku obserwacji naukowej, obowiązkowo należy opracować. Zgodnie z założonymi celami badacz powinien dokonać ich analizy, porównania, systematyzacji itd.

W praktyce przeprowadzania badań naukowo-pedagogicznych wykorzystywane są różne **rodzaje obserwacji**:

- obserwacja *bezpośrednia i pośrednia*. Pierwszą charakteryzuje bezpośrednie prześledzenie przebiegu obserwowanych procesów. Natomiast pośrednia obserwacja odbywa się wtedy, gdy sam proces jest ukryty przed obserwatorem (na przykład poznawcza aktywność uczniów), ale jego stan utrwała się na podstawie możliwości do obserwacji i opisu wskaźników (tempo opanowywania i objętość przyswajanej informacji, liczba i jakość zadawanych pytań, liczba i poziom rozwiązanych i nierozwiązanych sytuacji problemowych, jakość i szybkość ich rozwiązywania, stosunek uczniów do trudności intelektualnych itd.);
- obserwacja *ciągła i wybiórcza*. Pierwsza obejmuje cały proces w jego jednolitości od początku do końca (na przykład obserwacja lekcji). Druga polega na wybiórczym utrwalaniu badanych zjawisk (na przykład zachowywanie się dzieci w sytuacjach konfliktowych);
- obserwacja *otwarta i tajna*. W pierwszym wypadku badany wie o tym, że jest kontrolowany, w drugim – nie;
- obserwacja *włączona i niewłączona*. Zależy od pozycji obserwatora w stosunku do badanego. Jeżeli badacz jest włączony we wspólną działalność z badanym i obserwuje go niejako od wewnątrz, taką obserwację nazywamy włączoną, natomiast jeżeli obserwator znajduje się na zewnątrz obserwowanego procesu, obserwacja taka będzie niewłączona. Na przykład, w grze w siatkówkę przy włączonej obserwacji badacz jest członkiem drużyny siatkarzy, przy niewłączonej – kibicem obserwującym grę i graczy;
- obserwacja *krótkotrwała i długoterminowa*. Pierwsza zachodzi, gdy obserwuje się lekcję lub przedsięwzięcie wychowawcze, druga, gdy obserwuje się ucznia w ciągu roku szkolnego;
- obserwacja *longitudinalna* (podłużna) i *retrospektywna* (zwrócona w przeszłość). Na przykład podczas badania warunków sprzyjających rozwojowi zdolności matematycznych ucznia, nauczyciel obserwuje ucznia od pierwszej do ostatniej klasy (obserwacja longitudinalna). Widząc wybitne zdolności ucznia ostatniej klasy, pedagog bada jego dane biograficzne, opinie, wyniki nauczania, zainteresowania itd. w klasach poprzednich, wyodrębniając przy tym główną determinantę jego zdolności (obserwacja retrospektywna).

Aby uzyskać wiarygodną informację o zjawisku pedagogicznym, obserwacje powinny być wystarczająco długotrwałe, systematyczne, wielostronne, mo liwie obiektywne (przeprowadzone bez nadmiernego zainteresowania zarówno w rezultatach pozytywnych, jak i negatywnych), masowe, eby przy obróbce otrzymanych danych mo na było mówić, je eli nie o prawidłowościach, to chocia by o ujawnionych tendencjach w rozwoju badanych zjawisk.

Mimo bogactwa informacji, łatwości i du ego rozpowszechnienia metody obserwacji, pozwala ona ukazać wyłącznie zewnętrzne przejawy zjawisk pedagogicznych, natomiast najbardziej istotne procesy wewnętrzne, pozostają poza polem widzenia badacza. W związku z powy szym mo na stwierdzić, e otrzymany za pomocą ró norodnych obserwacji materiał ma własną wartość naukową, równocześnie wymaga on gruntownej obróbki i mo e stanowić podstawę dla zastosowania innych metod badań naukowo-pedagogicznych.

2. Metody wywiadu w pedagogice, wykorzystywane są w celu wyjaśnienia subiektywnego stosunku badanych do ró nych dziedzin działalności i wiedzy, a tak e do określenia poziomu opanowania wiedzy. Metody te mają formę pytanie – odpowiedź i przeprowadzane są ustnie lub pisemnie. Jest to **dialog badacza z badanym prowadzony według wcześniej opracowanego programu.**

Badanie ustne przeprowadza się podczas *bezpośredniego kontaktu* badacza z badanym. Takie badania mają dosyć elastyczny i szybki charakter, pozwalają uściślać, poszerzać, a nawet modyfikować badanie w trakcie jego przeprowadzania. Jednak podczas badania ustnego powstaje problem z rejestracją i obróbką jego rezultatów, a sam proces badania wprost zale y od komunikatywności eksperymentatora, co wpływa na poziom subiektywności badania.

Wśród metod badania ustnego nale y wyodrębnić **rozmowę i wywiad**. Obydwie formy są *aktem komunikacyjnym, w trakcie którego, otrzymując odpowiedzi na postawione pytania, badacz wyjaśnia istotę interesującego go zjawiska*. Ró nica między rozmową a wywiadem polega na wyraźnym rozgraniczeniu ról (pytający – odpowiadający) w trakcie wywiadu oraz bardziej swobodnej wymianie poglądów podczas rozmowy.

W uzyskaniu wiarygodnej informacji podczas badania ustnego wa ną rolę odgrywają konsekwencja i jakość stawianych pytań. Powinny one być sformułowane zwięźle, zrozumiale, jednoznacznie. W rozmowie badacza i badanego nie mogą znaleźć się pytania prowadzące w ślepe zaułki, pytania retoryczne, a tak e takie, które zawierają w sobie ocenę osobowości respon-

denta lub jego zdolności. W celu usunięcia przypadkowych odpowiedzi praktykowane są pytania uściślające i krzyżujące, potwierdzające lub obalające wypowiedziane wcześniej poglądy.

Liczba pytań w wywiadzie i rozmowie nie jest określona. Dialog trwa do czasu ujawnienia istoty badanego zjawiska, jednak nie wolno nadużywać cierpliwości badanych, ponieważ może to wywołać reakcję negatywną, niechęć do udziału w eksperymencie, a nawet podawanie odpowiedzi jawnie nieprawdziwych.

Technologia przeprowadzenia badania ustnego zakłada dobór kompetentnych respondentów, uzasadnienie i poinformowanie w dostępnej dla nich formie motywów przeprowadzenia badania. Dialog z respondentami powinien być etyczny, należy go przeprowadzać w poufnej i przyjaznej atmosferze, najlepiej w obecności badacza neutralnego. Eksperymentator nie ma prawa w żadnej formie narzucać badanemu własnej (lub upragnionej) pozycji.

Odpowiedzi respondentów należy obowiązkowo utrzymywać, jednak zaleca się nie czynić tego w obecności badanych. Lepiej je odtworzyć po zakończeniu badania. Choć z drugiej strony takie postępowanie niesie za sobą niebezpieczeństwo subiektywizacji wyników badania. Żeby go uniknąć, zaleca się wykorzystywać otwarte (o których wiedzą badani) i skryte formy zapisu.

Badanie pisemne opiera się na *pośrednim kontakcie* eksperymentatora z badanym. Jest ono realizowane za pośrednictwem ankiet (**ankietowanie**) i testów (**testowanie**). Jego wykorzystanie niweluje wpływ na badanie czynnika subiektywnego i podnosi wiarygodność obróbki jego wyników, chociaż jednocześnie obniża poziom procesowej elastyczności badania.

Do badania pisemnego stosuje się takie same wymagania jak do badania ustnego, z wyjątkiem kilku specyficznych czynników (liczba pytań w ankietach i testach jest ściśle ustalona, nie powinno być ich zbyt wiele, pytania w swojej formule nie powinny zawierać podpowiedzi). Do pisemnych kwestionariuszy w momencie przeprowadzenia badania nie wolno niczego dodawać lub dołączać. Dlatego, zanim będą one zastosowane na skalę masową, należy je poprzednio zaaprobować, co można uczynić poprzez przeprowadzenie tzw. programu pilotażowego, tj. badania próbnego na niewielkiej grupie badanych. Po czym, w razie konieczności, można dopracować formułę pytań, powielić materiały i przeprowadzić badanie właściwe.

Technika obróbki otrzymanych rezultatów zależy od liczby uczestników i stopnia złożoności zawartości kwestionariuszy. Obróbkę „ręczną” można wykonać poprzez wypełnienie tablic (matryc) i obliczenie typów odpowiedzi według wybranych kategorii. Obróbka maszynowa jest możliwa przy odpowiedziach indeksowanych i dających się sformalizować.

Istnieje kilka **rodzajów ankietowania**. Według sposobu przeprowadzenia można wyodrębnić ankietowanie:

- *kontaktowe*, gdy wywiad przeprowadza się przy bezpośrednim kontakcie eksperymentatora z badanym;
- *zaoczne*, przeprowadzane są za pośrednictwem kontaktów korespondencyjnych, gdy ankiety z instrukcjami rozsyłane są pocztą lub wręczane są osobiście osobom przeprowadzającym badanie i w ten sam sposób wracają na adres instytucji badawczej;
- *prasowe*, które realizuje się za pośrednictwem umiejscowienia ankiet w gazetach i czasopismach.

Ze względu na charakter przypuszczalnych odpowiedzi ankiety można podzielić na:

- *otwarte*, przewidujące każdą możliwą odpowiedź wybraną przez respondenta;
- *zamknięte*, przewidujące wybór jednej lub kilku odpowiedzi z szeregu zaproponowanych;
- *mieszane*, gdy badany może wybrać odpowiedź spośród już istniejących lub wpisać własny wariant odpowiedzi.

Szczególne miejsce w systemie metod badania pisemnego zajmuje **testowanie**. W czasach obowiązywania pedagogiki radzieckiej testy uważano za burzliwe i jako takie były zakazane w radzieckiej szkole i pedagogice. Pierwsza złamała ten zakaz psychologia, zastosowawszy testy w procesie określenia różnic indywidualnych i wykorzystania metody psychometrycznej. Później testowanie znalazło swoje miejsce również w pedagogice.

Obecnie testowanie psychologiczno-pedagogiczne stało się nieodłączną częścią teorii i praktyki pedagogicznej. Za pomocą testów przeprowadza się pomiar osiągnięć edukacyjnych, wyjaśnia się stopień przejawu cech osobowościowych, ujawnia się poziom rozwoju zdolności, stopień uformowania mechanizmów nauczania, wiedzy, kontaktów międzyludzkich itd.

Słowo „test” w tłumaczeniu z angielskiego oznacza „badanie”, „sprawdzian”, „próba”.

Testowanie należy rozumieć jako badanie ukierunkowane, jednakowe dla wszystkich badanych, przeprowadzane w ściśle określonych warunkach, pozwalające obiektywnie zmierzyć właściwości badanego procesu.

W zależności od dziedziny zastosowania, testowanie bywa fizyczne, intelektualne, psychologiczne itd. Proces testowania polega na wykonaniu przez badanego zadań, wymagających odpowiedniej działalności. Od innych sposobów badania, testowanie odróżnia się ścisłością, prostotą, łatwością, możliwością automatyzacji kontroli (Kozłowska 2001, s. 14–16).

Zazwyczaj w pedagogice wykorzystywane są testy intelektualne. Według charakteru przedstawienia zadań testy bywają:

- *proste*, gdy badany wie, co i po co jest badane (zazwyczaj są to testy szkolne, określające stopień przyswojenia materiału);
- *pośrednie*, gdy obiekt badania jest ukryty przed badanym (na przykład test kolorów Maksa Lüschera, w którym, na podstawie wybranej przez badanego gamy kolorów, ukazywane są jego cechy osobowościowe).

Testowanie jako metodę badań naukowo-pedagogicznych w praktycznej działalności pedagogicznej wykorzystuje się dla określenia poziomu bieżących wyników nauczania. Zazwyczaj w szkole stosuje się *testy końcowe*. Przeprowadza się je po ukończeniu większego rozdziału programu szkolnego i zawierają one około 120 (60) pytań przy 6-stopniowej skali ocen. Ocena jest wystawiana na podstawie liczby poprawnych (niepoprawnych) odpowiedzi. Do realizacji *bieżącej* kontroli wykorzystuje się zadania testowe. Zawierają one mniejszą liczbę pytań (6–12), dotyczących kluczowych zagadnień poznawanego materiału.

Warianty zadań testowych mogą być różne, np.:

- wybór poprawnej odpowiedzi spośród kilku zaproponowanych;
- wyłączenie z podanego zestawu odpowiedzi nieprawidłowej;
- rozmieszczenie wyszczególnionych jakości (cech) w poprawnej kolejności;
- zgrupowanie zjawisk według podanej cechy;
- wypełnienie luk w tekście;
- dobór synonimów lub antonimów;
- poprawa istniejących błędów itd.

W praktyce współczesnej szkoły często wykorzystywane są testy tematyczne o *róznicowanym poziomie*, dające możliwość zróznicowanego sprawdzenia stopnia opanowania materiału. Oprócz tradycyjnych, wykorzystywane są także testy *prędkości*, kiedy efektywność opanowania materiału wyjaśnia się na podstawie szybkości operowania terminami z wybranej dziedziny oraz *potęgi*, które stosuje się w celu określenia poziomu przyswojenia wiedzy na podstawie rozwiązania zadań złożonych.

3. Badanie dokumentacji szkolnej i wyników twórczej działalności uczniów jest źródłem obiektywnej informacji o procesie pedagogicznym i jego uczestnikach. Analiza pisemnych, graficznych, twórczych i innych prac uczniów daje badaczowi różnorodny materiał o motywacji i charakterze działalności poznawczej, o indywidualnych cechach dzieci, ich zaletach i brakach. Metoda ta, zawierająca różne treści, pomaga określić stosunek uczniów do przedmiotu, nauczyciela, kolegów oraz do innych, interesujących badacza zjawisk pedagogicznych.

ródnymi źródłami informacji badawczej są tak *e plany działalności szkoły* i jej części składowych, kółek, sekcji, *protokoły* posiedzeń organizacji społecznych, *materiały* rad pedagogicznych, księgi zarządzeń itd. Nieocenioną pomocą w pracy badawczej mogą okazać się *czasopisma szkolne, dzienniczki i zeszyty uczniów, akta osobowe dzieci, ich charakterystyki, formularze bibliotek szkolnych, karty medyczne* itd.

Uogólniając otrzymaną w ten sposób informację, badacz jest w stanie określić zależności, związki i stosunki empiryczne, które mogą posłużyć jako obiektywna podstawa dla kolejnych teoretycznych i praktycznych poszukiwań.

4. Metodę uogólnienia niezależnych opinii wykorzystuje się w celu podniesienia wiarygodności informacji o badanym zjawisku pedagogicznym. Cel ten można osiągnąć poprzez przeprowadzenie kilku niezależnych badań jednego przedmiotu, a następnie uogólniając otrzymane wyniki. Wskazaną metodę realizuje się na dwa sposoby:

- gdy jedno zjawisko pedagogiczne (na przykład, uczeń) jest badane przez kilku badaczy (nauczycieli, uczniów, rodziców) niezależnie od siebie, a otrzymane rezultaty są uogólniane, korygowane i uzupełniane,
- gdy jeden badacz bada jedno zjawisko, ale w różnych warunkach, stosując różne, niezależne od siebie metody, a następnie uogólnia otrzymane opinie.

Odmianą metody uogólnienia niezależnych opinii, a dokładniej, jej przedmiotem jest **konsylium pedagogiczne**. Metoda ta wymaga zespołowego udziału pedagogów w rozwiązywaniu powstałego problemu. Swoją nazwę metoda wywodzi z praktyki medycznej, ponieważ ma analogiczne cele i strukturę realizacji.

Etapy przeprowadzenia konsylium pedagogicznego:

- uogólnienie posiadanej wiedzy o badanym zjawisku;
- stwierdzenie nieprawidłowości i przyczyn ich powstawania (postawienie diagnozy);
- wytworzenie wspólnego programu działań, skierowanych na usunięcie ujawnionych braków (wyznaczenie terapii).

Konsylium pedagogiczne bywa *indywidualne*, gdy zachodzi potrzeba w zorientowaniu się w stanie rozwoju, edukacji i wychowania konkretnego dziecka oraz *grupowe*, gdy odpowiednie decyzje należy podjąć w stosunku do grupy dzieci (klasy, kółka, sekcji itd.). Konsylia bywają *planowe*, tj. wcześniej uwzględnione w planach pracy szkoły, grupy metodycznej, wychowawcy klasy itd. (tradycją stały się planowe konsylia grupowe odbywające się podczas promocji dzieci z jednej klasy do następnej oraz indywidualne – w przypadku planowanych wyróżnień uczniów) i *pozaplanowe* (przeprowadzane w miarę potrzeb powstających w trakcie procesu szkolno-wychowawczego).

Metoda badania monograficznego zakłada wielostronne i głębokie zbadanie jednego (mono) zjawiska pedagogicznego.

Na poziomie teoretycznym wyniki takich badań są uogólniane w monografiach, mających jednego lub kilku autorów. W ostatnim przypadku są one nazywane monografiami zespołowymi.

Na poziomie praktycznym zdanie: „monograficznie zbadano 30 dzieci z problemami”, oznacza, że było badane każde dziecko z osobna, a nie grupa trzydziścioro dzieci (w tym celu należałoby stworzyć taką grupę).

2.2. Badania teoretyczne w pedagogice

W założeniach metod teoretycznych badań naukowo-pedagogicznych leży badanie teoretycznych źródeł informacji pedagogicznej i teoretyczna obróbka rezultatów badań, przeprowadzonych za pomocą innych grup metod. Wykorzystuje się je do powtórnej obróbki posiadanych danych, a ich głównym celem jest poszukiwanie związków przyczynowo-skutkowych, ujawnienie tendencji i prawidłowości, budowa systemów, układanie programów itd.

Do **metod teoretycznych** badań naukowo-pedagogicznych zalicza się metody:

- *badania literatury przedmiotu;*
- *analizy konkretno-historycznej;*
- *przejścia od konkretności do abstrakcji i od abstrakcji do konkretności;*
- *modelowanie.*

1. Metoda badania literatury przedmiotu pozwala zapoznać się z wynikami już przeprowadzonych w danej dziedzinie badań, daje możliwość usystematyzowania materiału według wybranego problemu, znalezienia dziedzin zastosowania własnych wysiłków badawczych. W charakterze *materiału źródłowego* (bazy źródeł) przy wykorzystaniu tej metody występują dokumenty programowe (ustawy, decyzje, programy itd.), literatura monograficzna i informacyjna, zbiorowe prace naukowe, materiały z konferencji, periodyczne wydania pedagogiczne itd. Uwzględniając fakt, że obecnie najbardziej efektywne są poszukiwania w dziedzinach międzydyscyplinarnych, dotyczące dwóch lub kilku nauk, należy dołożyć wszelkich starań, aby uwzględnić ich osiągnięcia.

Użycie metody badania literatury przedmiotu wymaga od badacza opanowania całego szeregu *specjalnych umiejętności*, takich jak umiejętność korzystania z zasobów bibliotecznych, katalogów, informatorów bibliograficznych itd. Pracując z tekstem, badacz powinien umieć wyodrębnić kwestie zasad-

nicze, porównywać, uogólniać, systematyzować, w sposób uargumentowany udowadniać lub odrzucać znane i nieznanne fakty. Szczególnie w tej umiejętności jest opracowywanie w odpowiedni sposób zbadanego materiału: robienie wypisów, konspektów, stawianie tez, pisanie recenzji, adnotacji, referatów, artykułów naukowych, uzasadnianie przyszłych badań.

Oprócz źródeł poszukiwań teoretycznych literatura naukowa jest tak e wzorem przedstawienia i opracowania myśli naukowej.

2. Metoda analizy konkretno-historycznej opiera się na zasadzie społecznej zależności procesu pedagogicznego. Metoda ta zakłada badanie zjawisk pedagogicznych z uwzględnieniem miejsca, czasu i warunków ich przebiegu, a tak e związków i zależności, w jakich znajdowały się one lub znajdują się obecnie, tj. z uwzględnieniem warunków konkretno-historycznych i związków przyczynowo-skutkowych ich pojawienia się, przepływu i przekształcenia się. Doświadczenie historyczne pokazuje, że to samo zjawisko pedagogiczne (na przykład wariantowość szkolnych programów edukacyjnych) w jednych warunkach może być aktem postępowym, dającym miejsce wyborowi i twórczości zarówno nauczycielowi, jak i uczniom, a w innych warunkach może destabilizować system szkolny.

Umiejętne zastosowanie metody analizy konkretno-historycznej pozwala wybrać najlepsze osiągnięcia nauki pedagogicznej i unieść błędów popełnionych wcześniej.

3. Metoda przejścia od konkretno do abstrakcji i od abstrakcji do konkretno pozwala ukazać prawidłowości funkcjonowania systemów pedagogicznych i ich komponentów. W trakcie jej realizacji można wyodrębnić następujące etapy:

- badanie i opis zjawiska pedagogicznego;
- wyodrębnienie jego komponentów;
- zjednoczenie komponentów w całościowy system;
- zbadanie związków wewnętrznych i zewnętrznych systemu;
- ukazanie głównych cech badanego zjawiska;
- ujawnienie prawidłowości rozwoju i funkcjonowania badanego systemu.

U podstaw metody przejścia od konkretno do abstrakcji i od abstrakcji do konkretno leży *indukcyjna* droga poznania rzeczywistości pedagogicznej, kiedy badanie szczegółu doprowadza do odkrycia ogółu. Na przykład badanie działalności nauczyciela i uczniów na lekcji ukazuje prawidłowości przekazywania doświadczenia społecznego od jednego pokolenia do drugiego.

Włączając mechanizmy poznania *dedukcyjnego* (od ogółu do szczegółu), można mówić o metodzie **przejścia od konkretno do abstrakcji i od abstrakcji do konkretno**. Na przykład stworzywszy model nauczyciela twórcze-

go, mo na przejść do modelu twórczego nauczyciela muzyki, fizyki, języka obcego. W tym celu nale y wypełnić taki model odpowiednimi konkretami.

Metoda modelowania jest poglądowo-obrazowym ukazaniem pedagogicznych zjawisk, procesów, prawidłowości itd. w postaci *wzorów, schematów, tablic, matryc, wykresów, rysunków, rysunków technicznych, figur geometrycznych* itd.

Podstawowym atutem badacza korzystającego z tej metody, jest mo liwość umieszczenia w ograniczonej objętości wielu informacji o badanym zjawisku. Zwijanie i rozwijanie takiej informacji pozwala doskonalić działalność myślową, ukazuje pełną zawartość badanego procesu, pozwala znaleźć jego miejsce w szerszej wspólności zjawisk.

Mankamentem modelowania jest potencjalna mo liwość istnienia zbyt du ej liczby przypuszczeń, schematyzacji, odrzucenia szczegółów, co przy nieumiejętnym wykorzystaniu metody mo e doprowadzić do wyjałowienia samej teorii pedagogicznej.

2.3. Statystyczno-matematyczne metody badań pedagogicznych

Matematyzacja nauki uwalnia pedagogikę od jednostronnego, wyłącznie jakościowego, opisu zjawisk pedagogicznych, daje mo liwość wyjścia na poziom uogólniających abstrakcji i pozwala rewidować zdobycze nauki. W tym celu wykorzystywane są **metody statystyczno-matematyczne**, dające **narzędzia do ilościowego badania zjawisk pedagogicznych**. Wykorzystuje się je w dwóch celach: do obróbki danych, otrzymanych podczas wykorzystania innych metod oraz do diagnostyki, prognozowania i komputeryzacji zarówno procesu pedagogicznego, jak i badań naukowo-pedagogicznych.

1. Metoda udziału pozwala **stwierdzić udział (część, stopień, poziom) badanego zjawiska w procesie ogólnym**. Najczęściej do stwierdzenia stopnia takiego udziału wykorzystywana jest skala procentowa, bazująca na setnej części całości, rzadziej – skala poziomowa i stopniowa, pozwalająca zwłaszcza ukazać jakościowo-ilościowe kryteria-dystrybutory.

Metoda udziału pozwala przenieść charakterystyki pedagogiczne na język sformalizowany i daje mo liwość zestawienia nie tylko zjawisk podobnych, ale również wielkości zewnętrznie nieporównywalnych. Na przykład, lekcje literatury i wychowania fizycznego, muzyki i matematyki, efektywność działalności nauczyciela i uczniów.

2. Metoda rankingu pozwala ocenić określoną działalność przez kompetentnych sędziów według wybranej skali.

Zazwyczaj w celu oceny zjawisk pedagogicznych wykorzystuje się *skale interwałową*, w danym wypadku – *stopniową*. Może ona być *dwu-, trzy-, cztery-, pięcio-, ..., stu-stopniowa* itd. Każdemu stopniowi w wybranej skali odpowiada określona jakościowa charakterystyka badanego zjawiska. Dwa krańcowe punkty każdej skali oceniania zawsze są jednakowe. Pierwszy wskazuje na to, że badana właściwość nie występuje, a ostatni na to, że cecha przejawia się w pełnej mierze. Średnie położenia mówią o częściowym przejawianiu się badanego zjawiska.

Dwustopniowa skala zawiera dwie oceny:

- „0” – właściwość nie przejawia się (nie),
- „1” – właściwość przejawia się (tak);

Trzystopniowa skala:

- „0” – właściwość nie przejawia się (nie),
- „1” – przejawia się częściowo (czasem),
- „2” – przejawia się całkowicie (zawsze).

Czterostopniowa skala:

- „0” – ...
- „1” – raczej nie niż tak,
- „2” – raczej tak niż nie,
- „3” – ...

W badaniach pedagogicznych dość często stosowane są tak zwane *skale polarne*. Wykorzystuje się je dla jednoczesnego oceniania dwóch przeciwstawnych wielkości. Na przykład polarne skalowanie cech osobowościowych przy pięciostopniowej skali oceniania może być przedstawione w trzech wersjach:

| | | |
|----------------------------|---------------|--------------------|
| 1-a: zorganizowanie | +2 +1 0 -1 -2 | niezorganizowanie; |
| pilność | +2 +1 0 -1 -2 | lenistwo; |
| uczciwość | +2 +1 0 -1 -2 | kłamliwość; |
| 2-a: zorganizowanie | 4 3 2 1 0 | niezorganizowanie; |
| pilność | 4 3 2 1 0 | lenistwo; |
| uczciwość | 4 3 2 1 0 | kłamliwość; |
| 3-a: zorganizowanie | 0 1 2 3 4 | niezorganizowanie; |
| pilność | 0 1 2 3 4 | lenistwo; |
| uczciwość | 0 1 2 3 4 | kłamliwość. |

Potrzebną wersję z możliwych wybiera się w zależności od celów i szczegółów przeprowadzanego badania, ale zawsze znaczenie każdego stopnia wstępnie omawia się. Wskazana wersja zastosowania metody rankingu łatwo poddaje się obróbce maszynowej.

3. Metoda porządkowania polega na **konsekwentnym rozmieszczeniu badanych zjawisk według ich ważności**. W praktyce pedagogicznej wykorzystywane są dwa warianty tej metody: porządkowanie *według spadku*, występujące, gdy na pierwszym miejscu stawia się najbardziej znaczącą jakość, a na drugim – trochę mniej znaczącą itd., oraz porządkowanie *według wzrostu*, występujące, gdy na pierwszym miejscu znajduje się najmniej znacząca jakość, a na ostatnim – najbardziej znacząca.

Zazwyczaj w badaniach naukowo-pedagogicznych porządkowanie odbywa się według wskaźników ilościowych, otrzymanych w wyniku wykorzystania innych metod.

4. Metodę korelacji wykorzystuje się w celu **przedstawienia związków i zależności między dwoma badanymi zjawiskami**. Współczynnik korelacji (r), obliczony według określonych wzorów, wskazuje *stopień i charakter* tej zależności. Zmienia się on w granicach $[-1 \quad +1]$. Jeżeli wskazanie współczynnika korelacji będzie wahać się w przedziale $[0,7 \quad 1]$, to zależność między zjawiskami jest *istotna*, należy ją uwzględniać podczas organizacji procesu pedagogicznego. Natomiast jeżeli wskazanie współczynnika zmienia się w przedziale $[0 \quad 0,7]$ – to zależność *jest nieistotna*, można ją pominąć. *Ujemne* wskazanie współczynnika pokazuje, że istnieje odwrotnie proporcjonalny związek między badanymi zjawiskami, a *dodatnie* wskazuje na związek wprost proporcjonalny.

2.4. Zintegrowane metody badań pedagogicznych

Omówione powyżej grupy metod nie wyczerpują całej różnorodności narzędzi badawczych pedagogiki. W praktyce badań naukowo-pedagogicznych, oprócz stałego powiększania się liczby metod, pojawiają się także różne warianty metod zintegrowanych. Wzajemnie uzupełniając się i kontynuując, w konkretnym badaniu metody badań naukowo-pedagogicznych występują jako **zintegrowane**. Charakterystycznym ich przykładem są **metody eksperymentu pedagogicznego i metoda badania i rozpowszechniania nowatorskiego doświadczenia pedagogicznego**. Każda z tych metod zawiera szereg innych metod badań naukowo-pedagogicznych. Na przykład w skład eksperymentu konstatacyjnego wchodzi zarówno metody empiryczne (obserwacja, ankiety, badanie dokumentacji szkolnej itd.), teoretyczne (modelowanie, metoda przejścia od konkretnego do abstrakcji itd.), jak i metody statystyczno-matematyczne, wykorzystywane do obróbki otrzymanych rezultatów.

1. Eksperyment pedagogiczny polega na zorganizowanej w sposób specjalny pracy, mającej na celu ukazanie i rozwój określonych cech, zjawisk, procesów. Przeprowadza się go w warunkach rejestrowanych i ma on charakter twórczy. Od eksperymentu należy odróżnić pracę eksperymentalną. Jeżeli eksperyment skierowany jest na otrzymanie nowej wiedzy o zjawisku lub procesie pedagogicznym (dlatego zalicza się go do metod badań naukowo-pedagogicznych), to praca eksperymentalna polega na obróbce i doskonaleniu już posiadanej wiedzy. Na przykład praca eksperymentalna skierowana na wprowadzenie nowych programów szkolnych, którą najpierw prowadzi nowatorska grupa praktyków, a dopiero później włącza się główna grupa zainteresowanych osób.

Eksperyment skierowany na ukazanie i zbadanie określonych cech (zjawisk, procesów) nazywany jest **konstatującym**. Ma on wcześniej wyznaczony cel i wykonuje się go zgodnie z z góry opracowanym programem. Rezultaty eksperymentu mają wartość naukową wyłącznie w przypadku ich powtarzalności. Podczas eksperymentu konstatującego badacz znajduje „białe plamy” w badanym procesie lub zjawisku. Braki, które należy uzupełnić w trakcie eksperymentu **przekształcającego (formującego)**.

Analiza danych, otrzymanych w wyniku eksperymentu konstatującego i badania teoretycznego problemu, pozwalają sformułować hipotezę badania, tj. wysnuć *przypuszczenie naukowe, które podczas eksperymentu trzeba będzie udowodnić (lub obalić) poprzez doświadczenie*. Na podstawie hipotezy powstaje program i szczegółowy plan działań eksperymentatora, skierowany na rozwój pozytywnych cech lub zniwelowanie zjawisk negatywnych. Przed przystąpieniem do pracy eksperymentalnej opracowuje się aparat kategoryjny badania, tj. wyjaśnia się znaczenie wszystkich kategorii, którymi będzie operować badacz i uzasadnia się kryteria oceny badanego zjawiska.

Zazwyczaj dla przeprowadzenia eksperymentu pedagogicznego wybierane są dwie grupy badanych: *eksperymentalna i kontrolna*. W pierwszej realizuje się innowacyjne rozwiązanie problemu, w drugiej – tradycyjne. (Na przykład w jednej klasie nauczyciel wyklada swój przedmiot według programu eksperymentalnego, a w drugiej według ogólnieprzyjętego). Należy tak postępować w celu stwierdzenia efektywności pracy eksperymentalnej poprzez porównanie analogicznych wskaźników dla obu grup. Jednak są przypadki, gdy praca eksperymentalna przeprowadzana jest bez grupy kontrolnej (na przykład przy pracy z dziećmi stwarzającymi problemy); w takiej sytuacji porównywane są stany systemu w określonych odstępach czasowych i ukazywana jest dynamika zmian badanych zjawisk przed, w trakcie i po zakończeniu eksperymentu.

Według skali eksperymenty można podzielić na: **globalne**, tj. obejmujące dużą liczbę badanych (na przykład wprowadzenie nowych programów szkolnych), **lokalne**, obejmujące ograniczoną grupę uczestników (na przykład doskonalenie pracy wychowawczej w internatach) i **mikroeksperymenty**, przeprowadzane z minimalną liczbą uczestników (z uczniami jednej klasy lub nawet z poszczególnymi dziećmi).

Według charakteru działalności eksperymentalnej można wyodrębnić eksperymenty myślowe, statyczne i naturalne. Eksperyment **myślowy** polega na przeprowadzeniu działań i operacji eksperymentalnych w myślach. Dzięki wielokrotnemu odtwarzaniu sytuacji, często udaje się określić najbardziej korzystne warunki dla przeprowadzenia pracy eksperymentalnej, unikając oczywistych błędów, ominąć przeszkody. Eksperyment **statyczny** zakłada przeprowadzenie działań eksperymentalnych w warunkach laboratoryjnych, a **naturalny** – w warunkach naturalnych. Pierwszy z wymienionych eksperymentów jest analogiczny do planowania biznesowego, podczas którego stwarza się doświadczalny model, w celu jego sprawdzenia przed włączeniem do eksperymentu naturalnego, w którym badany znajduje się w realnym otoczeniu procesu pedagogicznego.

Według miejsca przeprowadzenia można wyodrębnić eksperyment naturalny i laboratoryjny. Eksperyment **naturalny** polega na naukowo zorganizowanym doświadczeniu mającym na celu sprawdzenie wysuniętej hipotezy; odbywa się on bez zakłócenia procesu edukacyjno-wychowawczego (w warunkach naturalnych). Taki rodzaj eksperymentu wykorzystywany jest w przypadku, gdy istota innowacji może przejawiać się tylko w warunkach rzeczywistych. Zazwyczaj obiektem eksperymentu są szkolne plany i programy, podręczniki i pomoce naukowe, formy i metody pracy edukacyjno-wychowawczej.

Jeżeli w celu otrzymania niezbędnych danych należy zapewnić szczególnie staranną obserwację badanego (na przykład z wykorzystaniem odpowiednich narzędzi), to praca eksperymentalna powinna odbywać się w specjalnie urządzonym pomieszczeniu – laboratorium. W takim wypadku eksperyment nazywamy **laboratoryjnym**. W badaniach pedagogicznych w związku z deficytem narzędzi metodycznych i laboratoryjnych stosuje się go rzadko.

2. Metoda badania i rozpowszechnienia nowatorskiego doświadczenia pedagogicznego. Nowatorskie doświadczenie pedagogiczne jest rzeczywistym źródłem teorii pedagogicznej, a mechanizm jego wprowadzenia w działalność współczesnej szkoły jest jednym ze sposobów nadania naukowego charakteru praktyce pedagogicznej. W metodzie tej znajduje swoje odzwierciedlenie aforyzm, mówiący o tym, że *teoria bez praktyki jest martwa, a praktyka bez teorii – ślepa i głucha*.

Metody badań pedagogicznych

Metoda badania i rozpowszechnienia nowatorskiego doświadczenia pedagogicznego polega na badaniu i teoretycznym uświadomieniu praktyki pracy nowatorskich szkół i nauczycieli. Poprzez doświadczenia nierzadko udaje się im rozwiązać kluczowe problemy pedagogiczne, które w teorii pedagogicznej często były przedstawione w sposób niewystarczający lub całkowicie pominięte. Je eli odkrycia pedagogiczne tego typu nie zostaną uogólnione, nie będzie określona ich baza teoretyczna i miejsce w teorii pedagogicznej, to pomimo technologicznego opracowania, nie uda się ich wprowadzić do praktyki działalności szkoły i nauczycieli. Pozostaną one nadal odkryciem dostępnym dla nielicznych.

Szczególne znaczenie nowatorskie doświadczenie pedagogiczne zdobyło w latach siedemdziesiątych XX wieku, w czasie krachu autorytatywnej szkoły i pedagogiki. Jako pierwsi wprowadzili nowy system edukacji nauczyciele-nowatorzy, później pojawiły się autorskie szkoły i metodyki, a następnie alternatywne systemy edukacji i wychowania.

Realizacja metody badania i rozpowszechnienia nowatorskiego doświadczenia pedagogicznego odbywa się według algorytmu zawierającego pięć podstawowych etapów:

1. Badanie praktyki pedagogicznej, wskazanie najbardziej efektywnych jej części.
2. Ustalenie zależności przyczynowo-skutkowych, determinujących wysokie osiągnięcia pedagogiczne.
3. Przedstawienie nowatorskich doświadczeń pedagogicznych w postaci systemu pedagogicznego, łączącego jego część treściową i technologiczną, teoretyczną i praktyczną.
4. Opracowanie materiałów metodycznych wdrażających nowatorskie idee i technologie pedagogiczne do praktyki edukacji i wychowania.
5. Rozpowszechnienie nowatorskich doświadczeń pedagogicznych: wydanie specjalistycznej literatury (książki, broszury, foldery itd.), przygotowanie materiałów audiowizualnych o autorach nowatorskich doświadczeń pedagogicznych i wynikach ich pracy, organizacja punktów konsultacyjnych, organizacja wystaw poświęconych nowatorskim doświadczeniom pedagogicznym itd.

Rozdział III

Miejsce metaanalizy w badaniach pedagogicznych

Pedagogika obejmuje swym zasięgiem myśl teoretyczną i dociekania empiryczne w badaniu zjawisk wychowawczych oraz wyjaśnianie istoty i zasad wychowania. Poznawanie jest zawsze z jednej strony systematyczne i wszechstronne, z drugiej – parcjalne i ograniczone. Pedagogika ma zatem jednocześnie naturę nauki empirycznej i teoretycznej. Właściwie mo na przyjąć, e funkcjonuje ona jedynie jako nierozzerwalny związek teorii z praktyką. Teoretyczne i empiryczne dociekania wzajemnie się uzupełniają. Empiria umo liwia pedagogice bezpośredni kontakt z rzeczywistością społeczną, przede wszystkim z rzeczywistością wychowawczą. Bez empirii pedagogika pozostawałaby na poziomie spekulacji i abstrakcyjnego myślenia o zjawiskach, a tak e bez treści uzyskanych poprzez kontakty z rzeczywistością wychowawczą. Brak myślenia teoretycznego i punktów wyjścia w pedagogice oznacza utratę mo liwości uogólniania i wyjaśniania zjawisk (nadawania im sensu). Uniemo liwia to również określanie istoty zjawisk i formułowanie praw ogólnych.

Jednym ze skutków zło oności zjawisk wychowawczych jest zło oność procesu poznawania istoty tych zjawisk. Oznacza to jednocześnie zło oność metod poznawania, czyli metod badawczych.

Szerzenie badań empirycznych słu y w znacznym stopniu rozwojowi metodologii pedagogicznej. Efektywność metody mo na ocenić i weryfikować głównie w empirycznej praktyce badawczej. Jej rozwój inicjuje dyskurs metodologiczny. Omówiona zale ność przebiega dwukierunkowo, czego rezultatem jest rozwój metodologii badań pedagogicznych i efektywność praktyki badawczej. Badania empiryczne stanowią podstawę rozwa ń teoretycznych.

Jednym z wa niejszych aspektów badań naukowych, przebiegającym w punkcie przejścia od empirycznych do teoretycznych dociekań o istocie zjawisk wychowawczych, jest problem integracji rezultatów badań empirycz-

nych. Integracja ta stanowi pomost pomiędzy faktami specyficznymi i już rozpoznanymi a prawami o charakterze bardziej ogólnym. Nie jest ona nowym problemem w metodologii naukowej. Jednak dylematy integracji, mimo ogromu prowadzonych badań do dziś pozostają ciągle otwarte. Nadal istnieje potrzeba poszukiwania nowych odpowiedzi na postawione pytania i otwierania kolejnych dróg w integracji badań empirycznych.

3.1. Rozwój badań metaanalitycznych

Każde dobre pierwotne badanie empiryczne zawiera we wczesnych fazach przegląd dotychczasowych teoretycznych doniesień naukowych. Jest on w swojej naturze próbą początkowej integracji. Przeglądy te w większości są niesystematyczne i opierają się głównie na intuicji. Z tych powodów opisy przeglądów rzadko posiadają naukowy charakter, o którym decydują między innymi takie cechy, jak: precyzyjność, rzetelność i jednoznaczność.

Założeń omówionych wyżej przeglądów jest sformułowanie przesłanek i hipotez badawczych. Analiza i synteza w takim ujęciu problemu pozostają zamierzeniami drugoplanowymi. Dlatego te przeglądy te w większości przypadków dobrze spełniają swoją funkcję. Analiza pierwotna w kolejnych fazach poszukuje odpowiedzi na pytania badawcze i przynosi nowe stwierdzenia teoretyczne. W przypadkach, w których dąży się do otrzymania z analiz pierwotnych nowych teoretycznych wniosków, a tym samym do zgłębienia istoty problemu badawczego, przeglądy te stają się niewystarczające. Można więc uznać, że mogą one okazać się wystarczające, ale do nowych wniosków nie prowadzą w sposób systematyczny. Stąd należy wyodrębnić przegląd jako wstęp do pierwotnego badania empirycznego i przegląd jako samodzielne badawcze przedsięwzięcie. Przeglądy te odróżniają zasadniczy cel oraz odmienna metodologia.

Najprostszym sposobem integracji ilościowej jest przeliczanie wyników potwierdzających lub obalających założoną hipotezę. W przypadku badań prowadzonych na próbach można przeliczać jedynie wyniki, które posiadają pewien poziom istotności (np. 0,05 lub wyższy). W omówionych przypadkach dopuszcza się nawet czynności, w których rezultaty posiadające poziom istotności poniżej 0,50 kategoryzuje się jako „pozytywne”, a rezultaty z poziomem istotności powyżej 0,50 kategoryzuje się jako „negatywne”. Często wyodrębnia się również trzecią kategorię – rezultaty graniczne, czyli średnie, nierozstrzygające wyniku w sposób zdecydowany (np. posiadające poziom istotności pomiędzy 0,40 a 0,60). Badanie z przeliczaniem „pozy-

tywnych” i „negatywnych” rezultatów przeprowadził Meehl w obszarze psychologii klinicznej (Meehl 1954). W swoim badaniu integrował rezultaty dwudziestu analiz pierwotnych, w których porównano prognozy przewidziane przez psychologów klinicznych i prognozy statystyczne. Meehl stwierdził, że prognozy statystyczne potwierdzały się w połowie badanych analiz pierwotnych, tylko w jednej analizie prawdziwsze okazały się prognozy psychologów klinicznych, a w pozostałych nie dowiedziono różnic. Nierównomierność wyników była tak oczywista, że Meehl tylko na jej podstawie wskazał na ułomność klinicznych prognoz. Inne badanie przyniosło zbliżone wyniki – podobne w sensie potwierdzenia badanych zależności, choć zupełnie innych. Dotyczyło ono efektów uczenia się z wykorzystaniem telewizji (Chu i Schramm 1968). Jego autorzy integrowali rezultaty badań skuteczności dwóch form uczenia się: klasycznego nauczania szkolnego i uczenia się za pomocą telewizji. Analiza dowiodła, że w 15% przypadków uczniowie nauczyli się więcej, korzystając z telewizji, w 12% skuteczniejsza okazała się nauka w szkole, w 73% przypadków nie dostrzeżono wyraźnej różnicy. Wyniki integracji pozwalają wnioskować przynajmniej o tym, że nie należy na podstawie poszczególnego badania jednoznacznie dowodzić o przewadze skuteczności jednej z form uczenia się.

Metodologia integracji badań, w których rezultaty wyrażone są w procentach jest bardziej skomplikowana. Przelicza się w nich wyniki potwierdzające lub obalające przyjęte hipotezy oraz kwantyfikuje się wkład każdego badania przy potwierdzaniu bądź obalaniu założonych hipotez. Przykład stanowią wyniki: 55% i 82%, z których pierwszy ma zdecydowanie mniejszy wkład, choć wspólnie leżą po tej samej stronie granicy, potwierdzając tę samą hipotezę.

Badanie Eysencka dotyczące efektywności psychoterapii jest przykładem zastosowania tej metodologii (Eysenck 1952). Objął nim szereg pojedynczych badań, które opisywały przemiany stanu neurotycznych pacjentów po odbytej psychoanalizie. Autor wyodrębnił z nich stopień trwałej poprawy, porównując go ze stopniem poprawy stanu pacjentów poddanych eklektycznemu leczeniu oraz pacjentów będących jedynie pod opieką lekarzy ogólnych lub pracowników socjalnych. Stwierdził on, że wśród pacjentów leczonych za pomocą psychoanalizy poprawa stanu zdrowia sięgała 44%, pacjenci poddani eklektycznemu leczeniu wykazali 64% poprawy, uzyskany wynik w trzeciej grupie to 72%. Oczywisty jest więc negatywny związek pomiędzy skutecznością leczenia i psychoterapią.

Sytuacja, w której wynikiem poszczególnych badań są wskaźniki korelacji, jest dla integracji bardzo prosta. Wskaźniki korelacji są same w sobie

miarą efektu i są one niezależne od jednostek pomiaru. W efekcie tego wskaźniki korelacji należą do najlepszych w procesie badania metaanalizy. Omówioną metodologię badawczą najlepiej ilustruje badanie Blooma, dotyczące wpływu różnorodnych czynników na osiągnięcia szkolne uczniów (Bloom 1976). Integruje się w niej wyniki badań osiągnięć szkolnych uczniów z 30 krajów. Badania te prowadzone były głównie w ramach International Studies of Educational Achievement (tabela A w aneksach). Dotyczyły one różnorodnych poziomów systemu szkolnego, różnorodnych klas w szkołach i różnorodnych przedmiotów nauczania. Bloom skoncentrował się na ocenie statystycznej wskaźnika korelacji. Interpretacja Blooma dotyczy przede wszystkim integrowanych rezultatów. Bloom do opracowania wskaźników korelacji nie wykorzystał metod statystycznych, które większość autorów uważa za nieodłączną cechę metaanalizy. Nie stosując ich, osiągnął jednak prawie taki sam naukowy poziom swoich badań jak w przypadku ich wykorzystania. Stało się to możliwe dzięki cechom wskaźników korelacji, które same w sobie są już pewnym odzwierciedleniem mocy efektu.

Erlenmeyer-Kimling i Jarvik opublikowali w latach sześćdziesiątych przegląd, odbijający się do czasu w świecie nauki (1963). Badali oni wpływ dziedziczności i środowiska na rozwój inteligencji. Mimo że otrzymane przez nich rezultaty nie zostaną przedstawione, należy podkreślić, iż stopień wykorzystania informacji zawartych we wskaźnikach korelacji był jeszcze wyższy niż w przypadku omówionych badań Blooma.

W tym punkcie Kulik i Kulik dostrzegają granicę między przeglądami a metaanalizami. Badanie prowadzone przez Erlenmeyera-Kimlinga i Jarvika autorzy ci klasyfikują już jako metaanalizę. Uważają nawet, że temu i podobnym badaniom do wymaganego poziomu brakuje tylko nazwy – metaanaliza.

W swoich pierwszych metaanalizach Glass i jego współpracownicy przedstawili koncepcję integracji kwantytatywnej (Glass i Smith 1977, Glass i Smith 1979, Glass, Cahen, Smith, Filby 1982, itd.). Autorzy ci jako pierwsi dokonali tak kompleksowych integracji. Rozległość tych metaanaliz jest zdecydowanie większa niż rozległość dotychczasowych przeglądów kwantytatywnych. Nikt przed Smithem i Glassem nie wykazał się tak wielką erudycją w aplikacji metod statystycznych dla potrzeb integracji badań empirycznych. Nikt też z taką uwagą nie śledził czynników wpływających na wyniki badań. Glass i Smith byli innowatorami w zastosowaniu różnic średnich arytmetycznych jako miary mocy efektu w integracji badań eksperymentalnych. Główną zaletą takiej miary mocy jest jej niezależność od jednostek pomiaru. Fakt ten w zasadniczy sposób rozszerzył obszar, który może być objęty przeglądami

kwantytatywnymi i integracją. Wskazali oni, że zasób badań, który można wykorzystać w integracji, jest znacznie większy od dotychczas zakładanego. Trzecią nowością jest prezentacja sposobu kontrolowania wpływu zbioru cech badań na rezultaty. Dotychczasowe przeglądy kwantytatywne dokonywały kontroli jednej lub maksymalnie dwóch cech. Bywały również takie, w których kontrola ta nie została przeprowadzona. Kolejną nowością dotyczy metod analitycznych zastosowanych przez Glassa i Smitha. Są one bardzo postępowe i zasadniczo odróżniają się od metod spotykanych w dotychczasowych kwantytatywnych przeglądach. Rozwinęli oni cały zbiór równań regresji, wykorzystywanych do badań dotyczących stosunku między efektem terapii a rodzajem terapii, charakterystyką leczonych, sposobem pomiaru efektów itd.

3.2. Główne cechy metaanalizy

W ostatnich dziesięcioleciach szczególnie dotkliwy stał się problem masowości publikowania rezultatów badawczych. Liczba wyników badań ulega ciągłemu zwiększeniu, co w konsekwencji prowadzi do ograniczenia możliwości dokonywania ich przeglądów. Naukowcy podejmujący te problemy w obszarze swoich specjalności, starają się dotrzymać kroku pojawiającym się wyzwaniom rzeczywistości. Prace naukowo-badawcze przeznaczone są przecież także dla nauczycieli, wychowawców, dyrektorów, pracowników oświaty itp. Konieczna i pilna jest więc lepsza systematyzacja tej mnogości literatury.

W naukowej i specjalistycznej literaturze światowej istnieją przeglądy (ang.: *review*), które próbują streszczać najważniejsze naukowe osiągnięcia i wyniki o określonej problematyce. Ponad dwadzieścia lat temu Glass odkrył niedociągnięcia ówczesnych naukowych przeglądów. Stwierdził on, że autorzy raportów i przeglądów prace naukowe do swoich integracji wybierają w większości zupełnie przypadkowo. Rezultaty i wyniki ich opracowań są nieprecyzyjne i zbyt ogólne. Z powodu wymienionych błędów stwierdzenia tych przeglądów są mało rzetelne, niesystematyczne i często nieporównywalne. Ponadto ich autorzy zbyt mało czytelnie opisują swoją metodologię. Słaba przejrzystość przeglądów uniemożliwia użytkownikom dokonanie oceny adekwatności prezentowanych stwierdzeń. Większość przeglądów ma zbyt prosty model metodologiczny o małej nośności. Główną przyczynę tego stanu Glass upatrywał w słabo rozwiniętej metodologii integracji wyników publikowanych badań. Dowodził on, że „potrzebujemy metod, które będą umożliwiały systematyczne wartościowanie wyników badań, aby można było

wyciągnąć istotę poznania z wielkiego zbioru poszczególnych badań” (Glass 1976, s. 4). Według Glassa przeglądy dokonywane ówczesnie nie mogą tego zadania zrealizować wystarczająco dobrze. Dlatego te wyodrębnił on trzy typy analiz niezbędnych do progresu prac badawczych w obszarze kształcenia i wychowania: pierwotna analiza, wtórna analiza i metaanaliza.

Pierwotna analiza stanowi podstawową czynność opracowania danych, którą zwykle wykonuje badacz, planujący badanie i zbieranie danych. Analiza wtórna jest reanalizą danych, ukierunkowaną na poszukiwanie odpowiedzi na pierwotne pytania badawcze za pomocą doskonalszych metod statystycznych. Metaanaliza to z kolei ilościowe opracowanie wyników badań, a nie danych pochodzących z pierwotnych i wtórnych analiz. Metaanalizy zatem dokonuje analiz statystycznych ilościowych wyników poszczególnych badań. Metaanaliza nie zajmuje się początkowymi danymi empirycznymi, lecz wynikami uzyskanymi z tych danych.

Dla metaanalizy charakterystyczne jest (Glass 1976, s. 3–4):

1. Wykorzystywanie najbardziej obiektywnych metod selekcji badań wybranych do syntezy;
2. Opisywanie cech wybranych badań z ilościowymi kategoriami (tzn. du a próba, mała próba; publikowanie w periodykach, monografie, itp.);
3. Wyniki wszystkich badań przekształcane są na jednakową skalę, zwaną „moc efektu” (ang.: *effect size*, ES);
4. U ywanie ilościowych metod statystycznych do pomiaru stosunku między cechami i wynikami badań.

Nie ka dy przegląd empiryczny jest metaanalizą. Dotyczy to głównie wcześniejszych przeglądów. Obok przedmiotu badań istotna jest również metodologia. Przedmiot badań zarówno w metaanalizie, jak i w przeglądzie mo e być taki sam, w przeciwieństwie do metodologii, która w obydwu przypadkach jest bardzo zróżnicowana.

Zamierzeniem autora jest dokonanie streszczenia głównych cech metaanalizy na podstawie dotychczasowych źródeł metodologicznych. Uwzględnione zostaną zasadnicze punkty wyjścia metaanalizy Glassa i prace powstałe w późniejszych latach (prace Slavina, Hedgesa, Rosenthala, Kulika i Kulik, McGawa, Huntera i Schmidta itd.).

Główne cechy metaanalizy mo na sformułować następująco:

1. Metaanaliza obejmuje wyniki badań (rezultaty z raportów badań). Oznacza to, e w badaniu metaanalizy nie zbiera się początkowego materiału empirycznego. Do metaanalizy potrzebne są badania pierwotne i niektóre ich rezultaty.

2. Metaanaliza jest aplikacją metod statystycznych na ilościowe rezultaty badań. Empirycznym materiałem metaanalizy są np. średnie, miary rozrzutu, wskaźniki korelacji, rezultaty weryfikacji hipotez statystycznych i inne.
3. Metaanaliza obejmuje zbiór poszczególnych badań. Niektóre z metaanaliz integrują kilkaset, a nawet kilka tysięcy badań.
4. W metaanalizie wylicza się moc efektu, a nie tylko jego kierunek lub ustalony poziom istotności. Procedura wyliczenia mocy efektu powinna umożliwić porównywanie efektów z poszczególnych badań.
5. Przedmiotem metaanalizy jest również stosunek wyników badań do ich zasadniczych cech. Uwzględnia ona również wpływ cech badań na ich wyniki. Celem metaanalizy jest nie tylko najprostsze streszczenie dostępnej literatury, ale także ustalenie wpływu cech badań na różnice, efekty, poziomy istotności lub na wskaźnik mocy efektu.

3.3. Koncepcje rozwoju metaanalizy

W pierwsze lata po wystąpieniu Glassa dostarczyły materiałów wskazujących jednoznacznie, że niektórzy autorzy problemem integracji badań empirycznych zajmowali się przed Glassem albo niezależnie w tym samym czasie. Stan ten doprowadził w efekcie do procesu porównań i integracji nowych metod. W rozwoju tym obok Glassa (i jego zespołu) fundamenty utworzyło jeszcze kilku metodologów.

Jednym z nich jest Rosenthal, który opracował taksonomię metod metaanalizy (Rosenthal 1991). Jednocześnie Hedges rozwinął nowoczesne statystyczne metody metaanalizy (Hedges 1984, Hedges i Olkin 1985). Interesujący jest także wkład Slavina, który rozwinął metodę najlepszych dowodów (Best Evidence Synthesis) w integracji badawczej (Slavin 1986). Istotnych podwalin metodologicznych dostarczyli Hunter i Schmidt, wskazując na nowe modele metaanalizy (Hunter, Schmidt i Jackson 1982) oraz na zastosowanie korelacji w metaanalizie (Hunter, Schmidt i Coggin 1988).

W czasie, gdy powstawały wymienione fundamentalne prace, badacze Kulik i Kulik przeprowadzili kilka metaanaliz, zaliczanych do grona najważniejszych w świecie (Kulik i Kulik 1982, 1985). W końcu lat osiemdziesiątych na podstawie swych doświadczeń badawczych napisali również kilka metodologicznych prac o metaanalizie (1986, 1989).

3.3.1. Badania Rosenthala

W początkach lat sześćdziesiątych badania własne doprowadziły Rosenthala do problemu kwantytatywnej syntezy badań empirycznych (Rosenthal 1963). W połowie lat siedemdziesiątych opublikował on rezultaty metaanalizy, którą objął 311 pojedynczych badań (Rosenthal 1976). Zastosował w niej własną metodę pomiaru mocy efektu wyprowadzoną z poszczególnych badań z zastosowaniem wskaźnika „d” – standaryzowanej różnicy średnich arytmetycznych między grupą eksperymentalną i kontrolną. Analizował statystycznie także stosunek pomiędzy cechami badań a wskaźnikiem „d”. Zastosowane metody nieustannie rozwijał, co doprowadziło w konsekwencji do opublikowania głównego metodologicznego dzieła o metaanalizie (Rosenthal 1984 i 1991).

Metody wykorzystywane w badaniach metaanalitycznych zostały przez Rosenthala sklasyfikowane na osiem kategorii, funkcjonujących w obrębie trzech zasadniczych kierunków:

- 1) syntetyzowanie lub porównywanie rezultatów;
- 2) obliczanie miar mocy efektu lub prawdopodobieństwa;
- 3) integrowanie dwóch lub większej liczby badań.

Idea Rosenthala postulowała rozwinięcie odrębnych metod statystycznych dla każdego z wymienionych kierunków.

Kolejną cechą metody Rosenthala jest ogromna tolerancja przy stosowaniu i ocenianiu rozmaitych metod oraz procedur statystycznych w obrębie metaanalizy. W swoich krytykach i analizach dokonywał on oceny różnorodnych metod i koncepcji metodologicznych. Porównywał często używaną metodę przeliczania pozytywnych i negatywnych rezultatów (rezultatów potwierdzających lub obalających określoną hipotezę) z metodą analizy wariancji Cochra. Wskazał na ogromną różnicę wyników tych metod z uwzględnieniem takich samych danych.

Jako jeden z niewielu badaczy w świecie Rosenthal dopuszcza możliwość integracji poziomów istotności nawet w przypadku dysponowania jedynie dwoma badaniami. Zakłada także możliwość zastosowania metaanalitycznych metod do porównywania rezultatów dwóch własnych eksperymentów. Najbardziej kontrowersyjnym punktem metodologii Rosenthala jest niewątpliwie ocena miary mocy efektu wyprowadzona z wartości testów statystycznych (np. Z, t itd.) i wielkości prób każdego pojedynczego badania. Wielu pozostałych metodologów (patrz: Kulik i Kulik 1985 i 1986) określa to ujęcie jako niewystarczające. Sądzi się, że przy ocenie mocy efektu należy uwzględniać np. model eksperymentu, w którym efekt został otrzymany. Istot-

ne jest również to, jak przebiegała kontrola sytuacji eksperymentalnej (np. kontrola blokowa, przyjęte techniki wyrównywania grup itd.). Tak samo problematyczne jest podejście Rosenthala, dotyczące oceniania mocy efektu z poziomów istotności i liczebności prób. Stwarza to jeszcze mniejsze możliwości rzetelnej oceny mocy. W przypadku stosowania poziomu istotności koniecznością staje się ustalenie rodzaju testu, określającego ten poziom. Nie jest obojętne również to, czy poziom istotności uzyskano w oparciu o parametryczne lub nieparametryczne testy. Dla tych samych poziomów istotności i liczebności prób, miara mocy efektu może zostać rozproszona w dużym przedziale, w zależności od zastosowanego testu. Różnym występującym problemem jest to, że bardzo często autorzy rozpraw dotyczących badań empirycznych referują jedynie poziom istotności, pomijając wyliczone parametry. Często też nie określa się rodzaju zastosowanego testu.

Badania takie stawiają metaanalizę przed dylematem: dokonać wyłączenia badań i zrobić metaanalizę jedynie z kompletnym raportem, czy zastosować badania z niekompletnymi raportami. Wyłączenie niektórych badań z metaanalizy zawsze oznacza zagubienie pewnych informacji. Dlatego też Rosenthal poszukiwał metody skutecznej w takich przypadkach, zdając sobie jednocześnie sprawę z ich niedociągnięć i niewielkiej mocy. Zmniejszanie zakresu lub zbyt wielka restrykcja wyboru badań do metaanalizy może oznaczać regresję i zwrot ku wcześniejszym modelom przeglądów kwantytatywnych. W takich okolicznościach doniesienia Rosenthala miałyby z pewnością większą szansę na akceptację niż potępienie. Główny problem dostrzega się w niebezpieczeństwie, polegającym na dokonaniu uchybienia typu „niedocenie błędu standardowego” z powodu zastosowanej metodologii. Takie uchybienia prowadzą w efekcie do dodania wyszeregowanego poziomu istotności ostatecznym rezultatom. Oznacza to jednocześnie niedocenie ryzyka odrzucenia statystycznej hipotezy zerowej. Efektem tego jest stronniczość przy rozwadnianiu, które z rezultatów obalają lub potwierdzają założoną hipotezę. Konsekwencją tego mogą być błędne uogólnienia i wnioski. Nie są to jednak wystarczające argumenty do całkowitego odrzucenia tego podejścia, lecz raczej do kolejnych rozważań teoretycznych dotyczących omawianego problemu.

Aby zapobiec mylnemu rozpoznaniu znaczenia wkładu Rosenthala w rozwój metaanalizy metodologii należy podkreślić, że zaprezentowano tu jedynie najciekawsze i najbardziej kontrowersyjne fragmenty jego koncepcji.

3.3.2. Restrykcja wyboru badań

Restrykcją wyboru badań do metaanalizy zajmował się również Slavin. Swoją metodologię rozwinął jako alternatywę wobec klasycznych przeglądów kwantytatywnych i wobec metaanalizy. Główną słabość metaanalizy dostrzega on w tym, że łączy ona bardzo różnorodne badania. Różnorodność ta dotyczy przede wszystkim poziomu zastosowanej metodologii. Zdaniem Slavina wszystkie rezultaty i wnioski płynące z badań, nie posiadają jednakowego ciężaru. Zasadniczą ideą jego metodologii jest integracja polegająca na statystycznej i racjonalnej analizie mniejszej ilości specjalnie wybranych badań, które autor integracji ocenia jako najistotniejsze dla wybranej tematyki, a jednocześnie wystarczająco poprawne pod względem metodologii.

Tak zaprezentowana metodologia jest według Slavina połączeniem zalet klasycznych kwantytatywnych przeglądów i metaanaliz. Z klasycznych przeglądów wybrano jako najistotniejsze dobor najlepszych badań, a w metaanalizie sięgnięto po kwantytatywne i statystyczne narzędzia. Metoda opracowana przez Slavina umożliwia badaczowi oparcie się na najistotniejszych badaniach, czyli na najistotniejszych wynikach badań. Określenie „metaanaliza” uznał za mało odpowiednie i opracowaną przez siebie metodę nazwał „syntezą najlepszych dowodów” (*best – evidence synthesis*).

Slavin swoją metodę zaprezentował także w sposób praktyczny. Lata osiemdziesiąte przyniosły kilka jego metaanaliz, dokładniej syntez najlepszych dowodów (Slavin 1987a, 1987b, Slavin i Karweit 1984 itd.). Z raportów tych metaanaliz jasno wynika, że empiryczny filar badań jest niezwykle wąski. Pierwsza metaanaliza dotyczyła 43 pojedynczych badań, druga – 18, a trzecia jedynie ośmiu. Pozostali autorzy śledzący koncepcje Slavina swoimi metaanalizami także obejmowali stosunkowo mało badań (np. Becker 1988). Należy jednak nadmienić, że dobor tych badań był wyjątkowo staranny.

Słabość omawianej metodologii polega głównie na tym, że kryteria doboru i mała liczba badań zmniejszają rozproszenie wykorzystanych rezultatów. Umożliwia to kwantyfikację stosunku między cechami badań i rezultatami. Kulik i Kulik oceniają, że to podejście prowadzi do spekulatywnych stwierdzeń (Kulik i Kulik 1989, s. 255). Autorzy ci analizowali kilka metaanaliz przeprowadzonych zgodnie z metodologią Slavina. Dowiedli oni jej kolejnych uchybień, polegających głównie na subiektywności badacza przy użyciu kryteriów wyboru analiz pierwotnych, a w efekcie wyboru stronniczego. Badacze ci twierdzą, że nawet wyjątkowo restrykcyjne kryteria nie zapewniają doboru najbardziej odpowiednich (najlepszych, najbardziej relewantnych) badań (Kulik

i Kulik 1987 i 1989). Oceny te budzą jednak wątpliwości, zwłaszcza w kwestiach zasadniczych rozważań nad metodologią Slavina. Należy przede wszystkim zauważyć, że u Slavina ukazał ogromne znaczenie metodologicznych kryteriów wyboru badań. Podejście to chroni przed trywialnymi wynikami i stwierdzeniami. Jest to szczególnie ważne w przypadkach, gdzie wyniki uzyskiwane są na bardzo wielkich próbach i ta właśnie wielkość jako jedyna decyduje o ich znaczeniu w końcowych ustaleniach metaanalizy. Koncepcja Slavina posiada potęgę niedociągnięcia w ujęciu, gdy metaanaliza postrzegana jest jako „nadanaliza”, dostarczająca ostatecznych i niepodważalnych odpowiedzi. Słabość ta nie wydaje się tak istotna, gdy metaanaliza nie próbuje negować częściowych twierdzeń płynących z poszczególnych badań.

3.3.3. Ocenianie mocy efektu

Trzeci z twórców teorii metaanalizy – Hedges w swoich pracach zajmował się przede wszystkim problemami oceniania miary mocy efektu oraz oceniania wariancji. Dwa wyróżnione przez niego parametry umożliwiają przeprowadzenie testów statystycznych hipotez zerowych. Autor w swoich opracowaniach nie poprzestał jedynie na ustaleniach Glassa, ale rozwinął proponowane przez niego metody statystyczne. Glass w swoich metaanalizach używał stosunkowo znanych i często stosowanych metod. W normalnym zastosowaniu metody te służą do opracowania wyjściowego materiału empirycznego. Glass natomiast wykorzystał je do opracowania **rezultatów** badań empirycznych. Hedges dostrzegł słabości i uchybienia użycia metod konwencjonalnych do omówienia niekonwencjonalnych danych. Niedoskonałości te stały się punktem wyjścia jego najważniejszego dzieła, pochodzącego z 1985 roku, zatytułowanego *Metody statystyczne dla metaanaliz* (Hedges i Olkin 1985).

Pierwszym etapem jego rozważań teoretycznych jest estymacja miary mocy efektu. Jako punkt wyjścia tych rozmyślań przyjął miarę proponowaną przez Glassa. Wskazał, że dotychczasowe procedury estymacji (procedury, których zaczął używać Glass, a pozostali autorzy je rozwinęli) prowadzą do stronniczych ocen (ocena obciążona, zwana również tendencyjną). Procedura ta stała się podstawą rozważań m.in. Cohena (1988) dotyczących miary mocy efektu. Hedges swoje ustalenia oparł na tej mierze. W celu likwidacji stronniczości ocen rozwinął on procedurę korekcji ocen. Praktyka badawcza w stosunkowo krótkim czasie dowiodła, że różnice między nieskorygowanymi i skorygowanymi ocenami są niewielkie, a korekcja nie przyniosła oczę-

kiwanych korzyści. Kulik i Kulik w swojej metaanalizie (Bangert-Drowns, Kulik i Kulik 1983) dokonali oceny miar mocy efektu w dwóch wersjach: z procedurą Hedgesa i z jej pominięciem. Korelacja między skorygowanymi i nieskorygowanymi ocenami wyniosła 0,99 (Kulik i Kulik 1989, s. 244). Z powodu tak dużej korelacji autorzy ci sądzą, że stosowanie korekty nie ma większego znaczenia.

Dalszy wkład naukowy Hedgesa pochodzi z obszaru testowania i weryfikacji hipotez, dotyczących wpływu cech badań na moc efektu. Do tego celu proponował on test homogeniczności. W przypadku gdy występujące rozproszenie jest większe od oczekiwanego, występuje konieczność ustalenia jego źródła. Do tego celu Hedges proponuje procedurę analogiczną do analizy wariancji, za pomocą której próbuje się ustalić, czy źródłem różnic są specyficzne cechy badań. Cel ten można osiągnąć poprzez rozdzielenie badań według określonej cechy na dwie lub więcej grup (np. publikowane i niepublikowane badania). Dalsza procedura dotyczy testowania stosunku między wariancją występującą między grupami i wariancją występującą w grupach. Można na ten sposób potwierdzić, czy określona cecha wpływa na moc efektu.

Hedges krytycznie oceniał klasyczne statystyczne metody stosowane w metaanalizie. Rozwinął on cały szereg nowych procedur statystycznych. Hedges obalił argumenty przemawiające za użyciem analizy wariancji w ramach metaanalizy. Oczywiście jest więc, że to właśnie jego wkład w rozwój metod statystycznych w metaanalizie jest największy. Niektórzy autorzy w ostatnich latach zwracają uwagę na konieczność zachowania ostrożności przy wykorzystywaniu niektórych metod proponowanych przez Hedgesa (McGaw 1988, Kulik i Kulik 1989). Najrzetelniejsi badacze tego problemu to Kulik i Kulik, którzy kontestowali, że metody Hedgesa muszą być poddane dokładnej i wnikliwej analizie (Kulik i Kulik 1989, s. 250–251). Pomimo że taka ocena brzmi zbyt ostro, propozycje dotyczące dokładnej i wnikliwej analizy mogą okazać się korzystne i niezwykle przydatne dla dalszego rozwoju metaanalizy metodologii.

3.3.4. Integracja metaanaliz

Synteza rezultatów metaanaliz jest ostatnim ogniwem w rozwoju metaanalizy. Charakteryzuje ją integracja i opracowanie wyników pojedynczych metaanaliz.

Jedno z takich badań przeprowadził Fraser i jego współpracownicy (Fraser 1987). Jest to wyjątkowo obszerne badanie dotyczące efektywności procesu

kształcenia. Autorzy określają problem badawczy jako badanie czynników wpływających na rezultaty kształcenia. Precyzyjniej można sformułować go jako identyfikację i oszacowanie kauzalnych stosunków między czynnikami początkowymi (wraz z występującymi w przebiegu procesu kształcenia czynnikami) a rezultatami stanowiącymi efekt kształcenia. Autorzy przez wstępne czynniki określają cechy uczniów, strukturę socjalną, cechy nauczycieli, treści kształcenia itp. Czynniki występujące w przebiegu kształcenia to klimat nauczania, organizacja procesu nauczania itp. Badanie stanowiące przykład syntezy obejmuje 134 publikowane raporty metaanaliz, zawierające rezultaty 7827 badań pierwotnych. Uczestniczyło w nich ponad pięć milionów uczniów, nauczycieli, dyrektorów szkół i innych pracowników oświaty.

Rezultaty pojedynczych metaanaliz integrowano, wyliczając współczynniki korelacji, wskazujące na zależność między wybranym czynnikiem a rezultatami kształcenia (tabela B w aneksach). Autorzy badania zgrupowali ponad 120 zmiennych w siedem większych czynników, którymi są: szkoła, nauczyciel, uczeń, środowisko rodzinne ucznia, organizacja procesu uczenia, metody nauczania, strategie nauczania. Wyniki tabeli wskazują na fakt, że aden grupowy lub pojedynczy wskaźnik korelacji nie przekracza 0,50. Z rozkładu współczynników korelacji wyprowadzono dowód o zło oności wpływu i jego całościowym oddziaływaniu. Trzy czynniki: szkoła, metody nauczania i środowisko rodzinne mają mniejszy wpływ na zmienną zależną, pozostałe cztery, tj.: uczeń, nauczyciel, organizacja procesu nauczania i strategie nauczania. Autorzy dokonali także obliczenia średniego współczynnika korelacji, obejmującego wszystkie czynniki, otrzymując wartość $r = 0,20$. Współczynniki korelacji, posiadające wartość 0,20 lub większą, interpretowano jako istotne w praktyce. Liczba omówionych współczynników korelacji wynosi 53. Przyjęcie takiej interpretacji oznacza, że praktykę wychowawczą należy poprawić przynajmniej w wymienionych 53 aspektach. Dotyczy to jedynie wpływu na syntetyczną zmienną zależną. W przypadku uwzględnienia poszczególnych zmiennych zależnych, ilość rozpatrywanych aspektów potęgę nie wzrasta. Oczywiście w przypadku rezultatów nauczania języka ojczystego są inne wpływy niż w przypadku rezultatów nauczania matematyki itd. Konstatując problem syntezy metaanaliz, należy wyrazić oczekiwanie, że narastanie ilości kolejnych integracji umożliwi dalsze prace nad prowadzeniem metaanaliz.

3.3.5. Krytyka metaanalizy

W krótkim czasie po pierwszej metaanalizie dotyczącej skuteczności psychoterapii, pojawiły się krytyki oraz zarzuty podważające założenia metodologiczne (Eysenck 1978, Mansfield i Busse 1977, Presby 1978). Nowe zarzuty pojawiły się również, gdy Glass i Smith opublikowali rezultaty metaanalizy, dotyczące wielkości klas szkolnych (Slavin 1984, Educational Research Service 1980).

Omawiana krytyka zawarta została w trzech zasadniczych punktach:

1. Dokonane metaanalizy wiążą z sobą sprzeczne zjawiska (w konkretnych przypadkach sprzeczne badania); nie istnieje możliwość otrzymania logicznych i naukowo uzasadnionych wniosków za pomocą integracji badań, które zawierają rozmaite czynności pomiaru, różnorodne zmienne i zupełnie różne badane populacje.
2. Przeprowadzone metaanalizy zbyt wiele uwagi poświęcają badaniom o niskim poziomie jakości.
3. Metaanalizy zbyt mocno opierają się na publikowanych rezultatach (rezultaty publikowane w zasadniczy sposób różnią się od niepublikowanych).

Pierwszy z zarzutów jest uzasadniony z punktu widzenia trudności, które pojawiają się w integracji bardzo odmiennych rezultatów. Istnieje wiele możliwości wyboru dróg rozwiązań metodologicznych, dotyczących badania określonej tematyki. Stąd też pojawia się wiele różnorodnych podejść w poszukiwaniu rozwiązań na następujące pytania: jak badać? jak przebyć drogę od postawienia hipotez do ich weryfikacji? jak otrzymać empiryczne dane o zjawiskach wychowawczych? itd. W pierwszej kolejności należy podkreślić, że we wcześniejszych przeglądach kwantytatywnych badacze również korzystali z bardzo różnorodnego materiału empirycznego. Nie używali oni jednak skomplikowanych metod statystycznych. Dlatego też problem sprzeczności rezultatów nie był tak bardzo widoczny. Integracja była prowadzona w sposób intuicyjny. Badacze dość często potwierdzali jedynie teoretycznie dobrze uzasadnione i uargumentowane stwierdzenia. W przeglądach i w metaanalizie nie można jednak ograniczyć się tylko do dociekań, które stanowią wielokrotną replikę określonego badania. Drugim argumentem przeciwko omawianej krytyce jest rozwój metaanalizy w czasie pojawienia się sformułowanych zarzutów (Mansfield i Busse 1977, Presby 1978). Okazało się, że istnieją drogi rozwiązań problemów i dylematów integracji bardzo różnorodnych rezultatów. Przejrzystym dowodem na taką możliwość jest procedura obliczania miary mocy efektu. Jednak pomimo ogromnego postępu w rozwoju metaanalizy ciągle otwarte pozostają pytania i wątpliwości, powiązane z precyzyjnymi metodami integracji rozmaitych rezultatów.

Miejsce metaanalizy w badaniach pedagogicznych

Krytyka dotycząca różnorodnego poziomu badań wybranych do metaanalizy jest raczej krytyką niektórych metaanaliz Glassa niż ogólną krytyką nowej metodologii. Należy stwierdzić, że istnieją możliwości uwzględnienia poziomu metodologicznego („jakości”) poszczególnych badań. Nie tylko Glass, ale również inni teoretycy i praktycy metaanalizy rozwinęli szereg czynności dotyczących kontroli oraz uwzględnienia poziomu metodologicznego poszczególnych badań (np. Hunter i Schmidt, McGaw, Kulik i Kulik).

Najbardziej uzasadniona jest krytyka właściwości metaanalizy polegającej na zbyt mocnym opieraniu się na publikowanych rezultatach. Nie istnieje możliwość otrzymania danych o wszystkich niepublikowanych badaniach. Nigdy te nie staną się one tak dostępne, jak badania publikowane. Dowiedziono empirycznie, że sam fakt publikowania ma wpływ na rezultaty. Stwierdzenia krytyków wskazujące na fakt, że istnieją różnice między rezultatami niepublikowanych i publikowanych badań są więc prawdziwe. Rodzi się zatem pytanie: czy istnieje konieczność rezygnacji z integracji rezultatów dostępnych z powodu braku uwzględnienia rezultatów niedostępnych, nawet w przypadku, gdy rezultaty te są niezwykle istotne? Krytykę tę należy rozumieć dwutorowo. W pierwszym sensie, że metaanalizy powinny dołożyć wszelkich starań, aby otrzymać także badania niepublikowane i ich rezultaty. Drugi sens wskazuje na potrzebę odnalezienia możliwości uwzględnienia wpływu publikowania na rezultaty już w trakcie dokonywanych czynności statystycznych (a nie tylko w ostatecznych interpretacjach). Metaanaliza powinna rozwinąć procedurę precyzyjnego uwzględnienia omawianego wpływu.



Rozdział IV

Źródła w integracji badań pedagogicznych

4.1. Definiowanie problemu badawczego

Zasadniczą odmiennością tradycyjnych przeglądów kwantytatywnych jest ograniczona możliwość integracji. Dlatego też w takich przeglądach dobór badań nie leży w centrum uwagi ani badacza, ani czytelnika przeglądu. Istotą przeglądów stanowi bowiem szeregowanie i opisywanie rezultatów badań. Integracja natomiast jest bardziej jakościowa niż kwantytatywna. Istnieje więc możliwość przypisywania przez badacza poszczególnym badaniom odmiennych znaczeń. Wszystkie procedury z tym związane są przypadkowe i bardzo niesystematyczne. Do celów przeglądów kwantytatywnych okazały się jednak wystarczające.

Prawidłowe rozpoznanie materiału badawczego, czyli właściwy dobór badań jest jednym z najważniejszych etapów badania metaanalitycznego. Etap ten charakteryzuje przede wszystkim bezkresna przestrzeń badawcza. Cecha ta powoduje w konsekwencji wykorzystywanie do metaanalizy nieograniczonej ilości badań.

Kulik i Kulik problemem doboru zajmowali się bardzo intensywnie, badając go teoretycznie i empirycznie. Z powodu potrzeb metaanalizy studiowali oni także dotychczasową praktykę integracji badań. Dowiedli oni, że w przeglądach kwantytatywnych dobór badań był podyktowany częściej przypadkiem niż cechami badań i ich metodologicznym modelem (Kulik i Kulik 1989, s. 257).

Możliwość metaanalizy pozwala dostrzec inny aspekt omawianego problemu. Dla metaanalizy nie istnieją granice ilości wykorzystanych badań. Pozwala to sądzić, że dobór badań stanowi uboczny problem metaanalizy. To odczucie jest jednak złudne, ponieważ biorąc pod uwagę to, że w metaanalizie analizuje się cechy badań, dobór ich należy do zadań zdecydowanie poważniejszych niż w przeglądach kwantytatywnych. Dobór badań ma aspek-

ty merytoryczne i metodologiczne. Pierwsze są ściśle związane z definiowaniem problemu badawczego. Aspekty metodologiczne dotyczą głównie ustalania kryteriów doboru.

Problem merytorycznego aspektu doboru zaprezentowano na podstawie analizy idei Sagadina, poruszających kwestie głównych faz nieeksperymentalnych badań empirycznych (Sagadin 1993, s. 13–21). Metaanalitik ju w początkowej fazie definiowania problemu zobligowany jest do poznania obszaru prowadzonego badania. Nie jest to jednak wystarczające do ostatecznego określenia problemu. Konieczne staje się również poznanie teoretycznej części badań i otrzymanie przeglądu dotyczącego dotychczasowych osiągnięć badań empirycznych w wybranym obszarze.

Po pierwszej fazie obejmującej początkową definicję problemu badawczego należy omówić działania drugiej fazy, czyli poznawanie teorii o zjawiskach, które poddane są dalszym rozważaniom. Krytyczna analiza teoretycznych ustaleń o wybranym problemie ma na celu jego przybliżenie we wszystkich aspektach oraz wskazanie teoretycznych punktów wyjścia. Powinna ona udzielić odpowiedzi na pytania, które umożliwiają kontynuowanie określania problemu badawczego.

Kolejna faza dotyczy przeglądu fragmentu rzeczywistości, który będzie podlegał metaanalizie. Ten etap najbardziej różni metaanalizę od innych badań. Określanie problemu badawczego jest w ogromnej mierze uzależnione od dotychczasowych badań empirycznych. Rzeczywistość badana w metaanalizie nie stanowi bezpośrednio praktyki wychowawczej. Jest to przyczyną wytyczenia nowych granic rzeczywistości objętej badaniami. Przegląd dotychczasowych badań empirycznych wyraźnie wskazuje na pełny zakres obszaru podlegającego badaniu. Stanowi to punkt wyjścia do podjęcia decyzji, jakie badania zostaną poddane metaanalizie. Konieczność późniejszego kategoryzowania cech u tych badań pojedynczych wymusza uwzględnienie konsultacji z praktykami (nauczyciele, wychowawcy, itd.) i przegląd praktyk, w których powstawały badania pierwotne. Drugim podstawowym aspektem przeglądu praktyki badawczej jest późniejsza selekcja rezultatów z poszczególnych badań. Oczywiście główne decyzje powiązane z selekcją rezultatów będą wykonane wówczas, gdy problem badawczy zostanie ostatecznie określony i rozpocznie się etap doboru materiału empirycznego do metaanalizy. W toku przyjętych czynności dąży się do tego, aby przegląd stanu badań empirycznych umożliwił ostatnią fazę definitywnego określenia problemu badawczego.

Zasadniczym celem i oczekiwanym przez badacza rezultatem tego etapu jest wyodrębnienie szczegółowych pytań badawczych, na które należy szu-

kać odpowiedzi w kolejnych etapach badania. Wraz z uszczegółowieniem problemu i postawieniem pytań badawczych przyjmuje się zało one hipotezy badawcze jako przewidziane odpowiedzi sformułowanych wcześniej pytań. Ostatecznie określony problem badawczy umożliwia wyodrębnienie jednoznacznych kryteriów służących do doboru badań. Są one uzależnione od sposobu określenia problemu badawczego w empirycznej literaturze badawczej. Kryteria muszą zapewnić adekwatną restrykcyjność. Zbyt ogólnie definiowana problematyka prowadzi do tematycznie mglistych i słabo określonych metaanaliz. Z kolei zbyt restrykcyjne kryteria mogą stanowić powód, który uniemożliwi badaczowi odnalezienie wystarczającej ilości badań. Rygorystyczne kryteria posiadają jeszcze jedną wadę. Jest nią zbyt małe rozproszenie cech poszczególnych badań. Metaanaliza powinna integrować rezultaty empiryczne będące podstawą wyprowadzania kryteriów. Wymaga się od niej nowych ustaleń teoretycznych. Przy założeniu, że metaanaliza potwierdza jedynie te hipotezy, które zostały już częściowo zweryfikowane w pojedynczych badaniach, jest ona metodą mało skuteczną. Stąd też postuluje się o empiryczny materiał metaanalizy, będący wystarczająco różnorodny i dostatecznie rozproszony. Rozproszenie to jednak nie może zagrozić tematycznej jednoznaczności i jasności problemu badawczego.

4.2. Kryteria doboru badań

Sprawa metodologicznych kryteriów doboru badań pozostaje ciągle otwarta. Rozwój metaanalizy z punktu widzenia kryteriów jakości modelu metodologicznego przebiega w dwóch zasadniczych kierunkach. Pierwszy z nich wyodrębnił Glass i jego uczniowie, którzy stoją na stanowisku większej tolerancji przy uwzględnianiu jakości metodologicznej (Glass, McGaw i Smith 1981, Hunter i Schmidt 1990). Według wymienionych badaczy metaanaliza może objąć jednocześnie swoim zasięgiem eksperymentalne i *quasi*-eksperymentalne badania lub eksperymenty z różnorodnymi modelami wewnętrznej kontroli sytuacji eksperymentalnej itp. Argumentem przemawiającym za słusznością takiego podejścia jest konieczność badania wpływu cech. W celu poszukiwania rzetelnych odpowiedzi na pytania badawcze metaanaliza musi opierać się na dostatecznie dużej bazie empirycznej. Jest to możliwe wówczas, gdy do metaanalizy włączone zostaną badania, których cechy są wystarczająco rozproszone.

Drugi z wymienionych kierunków polega na bardziej rygorystycznych wymaganiach metodologicznych kryteriów doboru badań do metaanalizy.

Najwybitniejszym przedstawicielem tego nurtu jest Slavin, twierdzący, że nie wszystkie badania tak samo nadają się do uwzględniania ich w metaanalizie. W procedurze doboru badań istotne jest przede wszystkim uwzględnienie modelu metodologicznego i wybór jedynie takich badań, które posiadają wystarczająco dużą wartość metodologiczną. Oznacza to dokonanie wyboru badań, które zawierają rezultaty uzyskane za pomocą metodologii gwarantującej ich wystarczająco dużą wiarygodność. Zaprezentowane podejście przy doborze badań pochodzi z postulatu, że poziom prawdziwości i rzetelności ostatecznych ustaleń zależy bardziej od rezultatów słabszych niż lepszych. Faktem jest, że w badaniach naukowych jednym z najistotniejszych kryteriów jest prawdziwość i rzetelność nowych ustaleń. Czy jednak poziom ostatecznych ustaleń jest rzeczywiście determinowany najslabszymi z wykorzystanych rezultatów? Trudno o istotny argument przemawiający za słusznością takiego podejścia. Jednocześnie nie można negować faktu, że wszystkie rezultaty wpływają na rzetelność i prawdziwość nowych ustaleń.

W sensownym związku dwóch tendencji dotyczących jakości badań dostrzega się rozwiązanie tego problemu. Należy uznać główną ideę koncepcji Slavina, dotyczącą odmiennych ładunków poszczególnych badań, pochodzących z ich cech metodologicznych. Z drugiej strony należy uwzględnić także postulaty Glassa o kontroli statystycznej i konieczności uwzględniania tych cech. Stąd też w metaanalizie istnieje potrzeba wykorzystania w roli zmiennych wszystkich cech metodologicznych, które są możliwe do kwantytatywnego opisu. W efekcie takich poczynań najlepsze badania pod względem metodologicznym posiadałyby największy ładunek, słabsze badania ładunek odpowiednio mniejszy a do badań najbardziej wątpliwych, czasami zupełnie wykluczonych z metaanalizy.

Dodatkowej odpowiedzi na tak postawiony dylemat dostarczyć mogą dotychczasowe doświadczenia, płynące z praktyki badawczej. Istotnym przedsięwzięciem w rozwazanym problemie wydaje się powtórna analiza omawiana przez McGawa (1988, s. 679). Autorzy tego badania Landman i Dawes dokonali analizy badań eksperymentalnych, dotyczących efektywności psychoterapii. Z całkowitej liczby 475 badań Landman i Dawes wybrali te, które ich zdaniem były najbardziej poprawne pod względem metodologicznym. Moc efektu otrzymana na tej podstawie okazała się w ogromnej mierze podobna do mocy efektu otrzymanej w badaniach całego zbioru. Umożliwiło to autorom potwierdzenie ustaleń Smitha i Glassa dotyczących kryteriów doboru badań do metaanalizy. Zdaniem McGawa jest to wystarczająco silny argument, potwierdzający podejście, w którym cechy metodologiczne (rozważane jako jakość) stanowią podstawę statystycznej analizy, nie służąc jed-

nocześnie jako kryteria doboru badań. Takie wywody trudno zaakceptować bez wątpliwości. Rezultaty analiz Landmana i Dawesa mo na interpretować jako potwierdzenie poprawności większej restrykcyjności doboru badań. Uwzględnienie całego zbioru badań, postulowane przez McGawa nie przyniosło zasadniczo odmiennych rezultatów. Oznacza to, e prezentowane podejście nie ma istotnego znaczenia w metaanalizie. Rezultaty te przemawiają przynajmniej częściowo za słusznością koncepcji Slavina.

W obliczu zaprezentowanych punktów widzenia pojawia się pytanie: w jaki sposób mierzyć wagę ka dego badania lub ka dego pojedynczego rezultatu? Najprościej problem ten rozwiązał Slavin. Odrzucił on wszystkie badania niespełniające zało onych kryteriów, określając tym samym ich wagę jako zerową. Natomiast wszystkim badaniom, które włączył do metaanalizy, nadał jednakową wagę. Podejście to jest jednak wystarczające jedynie do momentu, w którym nie wystąpi mo liwość bardziej precyzyjnego pomiaru wagi badań.

Jedyną mo liwość prawidłowego określania wartości poszczególnych badań stanowią ekspertyzy metodologów. One jednak są subiektywne. Istnieje więc potrzeba większej liczby ekspertów i ocen. Rosenthal (1991) proponuje podwójną procedurę oceniania metodologii badań, polegającą na dwukrotnym dokonaniu oceny przez tego samego eksperta. Pierwszą ocenę nale y wykonać na podstawie opisu metodologii ka dego raportu badania. Zdaniem autora istotne jest, e pierwsza ocena dokonuje się, zanim ekspert pozna rezultaty ocenianego badania. W drugiej ocenie juror powinien uwzględniać opis metodologii oraz interpretację rezultatów. Do takiego oceniania nale y przygotować szereg pytań, czyli kryteriów oceniania metodologii badań, zaczynając od formułowania pytań od najbardziej ogólnych do najbardziej szczegółowych. Na ka de pytanie oczekiwana jest odpowiedź stanowiąca ocenę. Końcowym jej rezultatem jest suma wszystkich ocen lub ich średnia. Oceny te nale y wykorzystać jako wagi przy określaniu jakości badań. Jeśli wybrane badanie będzie miało np. trzy razy większą ocenę, w analizach statystycznych nale y nadać mu trzykrotnie większą wagę. Oznacza to, e wkład ka dego badania do ostatecznych wniosków jest wprost proporcjonalny do jego metodologicznej jakości. Podejście to nale y traktować jako istotny krok w poszukiwaniu bardziej obiektywnych sposobów doboru badań.

4.3. Procedury poszukiwania badań

Treści zawarte w tej części opracowania koncentrują się wokół problemów opisu sposobów odnajdywania i selekcji badań oraz określania rezultatów, których wykorzystanie w pracach nad metaanalizą jest bezwzględnie konieczne. Obok procedur poszukiwania rezultatów zostaną tu również zaprezentowane mechanizmy opracowywania rezultatów oraz ich rzetelności. Czynność poszukiwania badań w metaanalizie nie różni się zasadniczo od poszukiwania badań w tradycyjnych przeglądach kwantytatywnych. W obydwu przypadkach zasadniczy cel stanowi odnalezienie i otrzymanie wszystkich rezultatów dotyczących określonej tematyki.

Poszukiwanie i analiza publikowanych abstraktów jest pierwszym etapem procesu gromadzenia badań, stanowiących źródła prac nad metaanalizą. W procesie tym należy jednak wyraźnie wykroczyć poza analizę jedynie abstraktów. Materiał zdobyty jedynie podczas analizy abstraktów uniemożliwia dokonanie precyzyjnych ustaleń, dotyczących mocy efektu poszczególnych badań. Konieczne staje się podjęcie kontroli statystycznej sposobów i rozmiarów, w jakich cechy badań wpływają na ich rezultaty. W metaanalizie proces analizowania abstraktów może wystarczyć tylko w incydentalnych przypadkach. Jedynie pierwsza faza zbierania informacji o badaniach, które należy wykorzystać w metaanalizie, może opierać się na abstraktach. Obecnie istnieje na świecie wiele rejestrów, które publikują streszczenia prowadzonych badań: Sociological Abstracts, Ulrich's International Periodicals Directory, Psychological Abstracts, Linguistics and Language Behavior Abstracts, International Bibliography of Periodical Directory itd. Wszystkie wymienione rejestry posiadają opracowanie komputerowe, umożliwiające łatwy dostęp. Dlatego też w większości przypadków faza zbierania informacji nie trwa długo. Otrzymane tą drogą abstrakty umożliwiają pierwszą orientację w obszarze poszukiwań. Najczęściej są to streszczenia książek i periodyków. Mimo że publikowane książki i czasopisma nie stanowią jedynego źródła metaanalizy poszukiwań, są one najczęściej najistotniejszym elementem i dostarczają zdecydowanej większości potrzebnego materiału badawczego. Na podstawie informacji otrzymanych z abstraktów należy rozpocząć zbieranie źródeł. Istotną czynnością metaanalizy jest kategoryzacja publikacji. Z jednej strony ułatwia ona systematyczne zbieranie źródeł, z drugiej strony umożliwia analizę statystyczną wpływu cech badań na ich rezultaty statystyczne.

Źródła można podzielić na pięć kategorii:

- 1) książki lub monografie;

Źródła w integracji badań pedagogicznych

- 2) czasopisma naukowe;
- 3) dysertacje doktorskie;
- 4) czasopisma specjalistyczne, biuletyny informacyjne, niepublikowane raporty badań (dostępne często w nielicznych egzemplarzach w bibliotekach instytucji naukowych);
- 5) pozostałe prace niepublikowane.

W związku z dokonaną kategoryzacją powstaje pytanie dotyczące rzetelności źródeł w badaniach metaanalizacyjnych. Rosenthal omawia rezultaty przeglądów kilku dokonanych metaanaliz (Rosenthal 1991). Zajmuje się on rzetelnością miary efektu w źródłach rozmaitego pochodzenia. Ponieważ te są to wyniki starszego pochodzenia, autorzy niniejszej książki uzupełnili badanie kwerendą najnowszych publikacji. Warto podkreślić, że Rosenthal wyodrębnił tylko cztery kategorie źródeł (dwie ostatnie kategorie są połączone).

Tabela 1. Rzetelność źródeł w badaniach metaanalizacyjnych

| Para źródeł | Ilość metaanaliz | Rzetelność |
|------------------------------------|------------------|------------|
| czasopisma : dysertacje | 24 | 0,88 |
| czasopisma : źródła niepublikowane | 19 | 0,69 |
| dysertacje : źródła niepublikowane | 17 | 0,83 |
| książki : czasopisma | 16 | 0,81 |
| książki : dysertacje | 13 | 0,92 |
| książki : źródła niepublikowane | 13 | 0,97 |

Źródło: Rosenthal 1991 i badania własne.

Większość wskaźników rzetelności jest wysoka, a jednocześnie obserwuje się między nimi duże różnice. Jakość dokonywanych metaanaliz (większość z nich dokonał Glass z współpracownikami) gwarantuje ogromną trafność otrzymanych rezultatów. W procesie poszukiwania najważniejszych odpowiedzi, istotne jest przeprowadzenie kolejnych empirycznych analiz i teoretyczne wytłumaczenie oraz uargumentowanie otrzymanych wskaźników.

Rezultaty zaprezentowane w tabeli wskazują jednak na tendencje, które są istotne już z dzisiejszego punktu widzenia rozwoju metaanalizy. Wyniki z tabeli można tak interpretować w świetle metodologii Slavina. Jeżeli wszystkie użyte źródła posiadają podobną rzetelność, trudno odnaleźć choć jeden argument zezwalający na wyprowadzanie wniosków metaanalizy jedynie na podstawie części dostępnego materiału badawczego lub części wykorzystanych źródeł. Pomimo tego Slavin nie wskazywał bezpośrednio na

u ycie źródeł, jego koncepcje pośrednio dotyczą ich wykorzystania. Przy doborze najistotniejszych badań (zakłada to metoda Slavina) nale y za ka - dym razem ustalić kryteria doboru. Slavin nie jest w tych poglądach odosobniony. Podobne postulaty pojawiają się te w koncepcjach innych badaczy problemu. Ró nice dotyczą jednak rodzaju przyjętych kryteriów. Zdaniem Slavina postulowane kryteria nale ą głównie do gatunku metodologicznych, w przeciwieństwie do ogólnie akceptowanych modeli metaanalizy z kryteriami tematycznymi, jako priorytetowymi. Wysoka rzetelność źródeł nie uzasadnia więc poglądów Slavina w całości. Nale y jednak nadmienić, e idee Slavina są w części uzasadnione. Dotyczy to zaobserwowanych ró nic w rzetelności u tych źródeł. Ró nice te nie są wprawdzie wielkie, lecz zasługują jednak na uwagę. Dowodzą one, e pochodzenie źródeł nie jest bez znaczenia.

Nie mo na pominąć równie kwestii dotyczącej miar mocy efektu, pochodzących z ró nych źródeł. Miary te mogą ró nić się między sobą pomimo wysokich wskaźników rzetelności. Podstawą ustalania poziomu rzetelności jest metoda korelacji. Współczynnik korelacji wskazuje na zgodność otrzymanych rezultatów, a nie na ich identyczność. Mimo sytuacji, w których określone źródło dostarcza systematycznie ni szych miar mocy efektu ni inne źródło, współczynnik korelacji mo e być wysoki. W sytuacjach, gdy stosunek między otrzymanymi miarami mocy efektu pozostaje niezmienny, wskazując jednocześnie na wysoką rzetelność, między miarami jakości mogą występować wyraźne ró nice. Fakt ten w powierzchownym ujęciu zupełnie obala rezultaty otrzymane przez Rosenthala. Rzeczywistość jest jednak inna. Ró nice w wartościach miar mocy efektu ujawniają się w późniejszej analizie wpływu cech badań na ich rezultaty. Analiza ta umo liwia uzupełnienie interpretacji omawianych ró nic.

Istnienie wyraźnych ró nic między miarami mocy efektu, mimo wysokiej rzetelności, ilustruje tabela z badań własnych. Tabela wskazuje na średnie wartości mocy efektu, otrzymane z artykułów zamieszczonych w czasopi smach w zale ności od tego czy pozostałe źródła były dostępne czy nie. Pierwsza kolumna zawiera moc efektu z badań, w których istniała mo liwość wykorzystania przynajmniej jednego dodatkowego źródła poza materiałem zasadniczym. Kolumna druga dotyczy wyników mocy efektu z badań, w których materiał zasadniczy stanowił jednocześnie jedyne dostępne źródło.

Źródła w integracji badań pedagogicznych

Tabela 2. Średnia moc efektu otrzymana z artykułów zamieszczonych w czasopismach

| Dodatkowe źródła | Dostęp | Brak dostępu |
|------------------|--------|--------------|
| książki | 0,51 | 0,68 |
| dysertacje | 0,49 | 0,57 |

ródło: badania własne.

Tabela wskazuje na wyraźne różnice pomiędzy mocą efektu w przypadku dostępu i braku dostępu pozostałych źródeł. Oznacza to jednocześnie, iż w metaanalizie istnieje konieczność dystansu i krytycznej analizy otrzymanych rezultatów w pojedynczych badaniach. Wskazane różnice należy uwzględnić przynajmniej w końcowej fazie interpretacji.

4.4. Przesiew wyników badań

W procedurze gromadzenia źródeł staje się konieczne wyraźne określenie i sformułowanie płaszczyzny rezultatów, które zostaną włączone do rozważań metaanalitycznych. Informacje o wyselekcjonowanych rezultatach okazują się niezwykle pomocne, jak dowodzą dotychczasowe ustalenia badawcze, podczas określania miar mocy efektu. Obok wymienionych rezultatów wymagane są również dane dotyczące testowania hipotez statystycznych.

W sytuacjach gdy badania pierwotne obejmują szersze problemy badawcze niż metaanaliza, należy dokonać selekcji rezultatów. Jest ona określona definiowaniem problemu badawczego metaanalizy. Rezultaty badań pierwotnych mogą być bardzo zróżnicowane, dlatego należy wybrać te, które w metaanalizie są naprawdę potrzebne. Przy takim doborze konieczny jest przegląd wszystkich badań oraz wgląd w ich charakter. Przegląd taki umożliwia określenie rodzaju potrzebnych rezultatów. W czynnościach dokonywania obliczeń i integracji miar mocy efektu należy uwzględnić następujące dane:

- dane, które wyrażają efekt pojedynczego badania (średnie arytmetyczne i wariancje wszystkich prób lub populacji ujętych w badaniu, procenty, otrzymane we wszystkich grupach oraz współczynniki korelacji);
- dane dotyczące liczebności prób;
- dane dotyczące wartości parametrów uzyskanych z testów hipotez statystycznych oraz poziomu istotności osiągniętego w każdym pojedynczym teście.

Źródła w integracji badań pedagogicznych

W czynnościach dokonywania analizy statystycznej różnorodność potrzebnych danych jest ogromna. Z jednej strony określa ją ogólny problem badawczy metaanalizy, z drugiej – wyniki przeglądu poszczególnych badań, wybranych do metaanalizy. Rosenthal prezentuje listę siedemnastu grup danych, dotyczących każdego badania. Została ona zestawiona po kilkudziesięcioletnich doświadczeniach prowadzenia badań metaanalitycznych (Rosenthal 1991, s. 42–43).

Tabela 3. Lista Rosenthala

| | |
|-----|---|
| 1. | Wyczerpujące dane bibliograficzne |
| 2. | Dane o autorze badania (z powodu późniejszych kontaktów i konsultacji) |
| 3. | Płeć osoby, która zbierała dane |
| 4. | Status osoby, która gromadziła dane (np. student, doktorant, pracownik uczelni itd.) |
| 5. | Stosunek między osobą zbierającą dane a wykonawcą metaanalizy (np. student – profesor, podwładny – pracodawca itd.) |
| 6. | Płeć osób w badanej próbie lub populacji |
| 7. | Dane dotyczące doboru próby |
| 8. | Płeć wykonawcy eksperymentu |
| 9. | Sposób doboru wykonawców eksperymentu |
| 10. | Stosunki interpersonalne między wykonawcami |
| 11. | Precyzyjne określenie zmiennych zależnych |
| 12. | Specyfika modelu badania |
| 13. | Charakterystyka grup kontrolnych |
| 14. | Opis kontroli prowadzenia eksperymentu (podwójna obserwacja, wideorejestracja eksperymentu itd.) |
| 15. | Dane dotyczące wszystkich zmiennych, które mogą wpływać na: – zmienne zależne (zmienne kontrolne), – moc efektu, – poziom istotności |
| 16. | Dane dotyczące komunikacji między wykonawcą a osobami poddanymi badaniu (np. sposób motywowania do udziału w badaniach) |
| 17. | Oczekiwane efekty |

ródło: Rosenthal 1991.

Zaprezentowane wymagania są niezmiernie skomplikowane w realizacji, zwłaszcza w sytuacjach, kiedy metaanalizą próbuje się objąć maksymalnie dużo o dostępnych badaniach z określonej problematyki. Zbieranie tak szczegółowych danych wymaga od badaczy badań pierwotnych doskonałej znajomości metodologii metaanalizy, jaką posiadają wykonawcy metaanaliz. Z całą

Źródła w integracji badań pedagogicznych

pewnością mo na jednak przyjąć, e postulowane oczekiwania dotyczące wiedzy metodologicznej są, niestety, nierealne. Mo na zatem uznać, e stanowią one model, który jest wskazówką dla ka dego metaanalityka. Jednocześnie nale y nadmienić, e jego całkowita realizacja jest mo liwa jedynie w przypadkach incydentalnych i zupełnie wyjątkowych okolicznościach. Okoliczności te rodzą pytanie: jak postępować z badaniami, które nie zawierają wszystkich wymienionych przez Rosenthala danych? Nie jest istotna jedynie zasada, mówiąca, e w badaniu naukowym nie uznaje się kompromisu. Z nutą humoru rozmiary opisywanego problemu mo na określić następująco: niewystarczająca jest czynność polegająca na kserowaniu listy Rosenthala i rozesłaniu jej po całym świecie wszystkim badaczom obszaru pedagogiki. Problem jest o wiele głębszy i zawiera jeden z trudniejszych dylematów metodologii metaanalizy. W ka dej konkretnej metaanalizie nale y ponownie określić relacje między oczekiwaniami badacza a mo liwościami wpływającymi z praktyki badawczej. Nale y jednak pamiętać, e nie wolno dopuścić do sytuacji, w której zamiar szczegółowego zbierania danych empirycznych spowoduje zawężenie, rozumiane jako pominięcie niektórych badań, ich rezultatów i wniosków. Ka de pojedyncze badanie pierwotne dostarcza materiału, stanowiącego próbę odpowiedzi na dociekania naukowe. Skutkiem braku uwzględnienia któregośkolwiek z badań są niekompletne i niepełne odpowiedzi na badawcze pytania metaanalizy.

Dlatego te nale y uwa nie prześledzić wszystkie konkretne okoliczności dotyczące dostępnego materiału empirycznego. W niektórych przypadkach koniecznością staje się rozszerzenie listy Rosenthala, w innych – jej zredukowanie.



Rozdział V

Opis cech badań

Istnieją dwa zasadnicze cele zbierania danych, dotyczących cech badań. Pierwszym z nich jest statystyczna analiza wpływu tych cech na rezultaty badań. Drugi cel dotyczy ukierunkowania praktyki badań empirycznych. Kontrola statystyczna i analiza cech badań stanowią konieczny element metaanalizy. Wymaga to realizacji dwóch zasadniczych warunków. Do pierwszego z nich zalicza się identyfikację i opis cech badań. Warunek drugi stanowi wystarczająco duże rozproszenie tych cech. Opisane i kategoryzowane cechy stanowią materiał empiryczny wykorzystywany w analizie statystycznej. Analiza powinna dowieść, w jaki sposób poszczególne cechy wpływają na miarę mocy efektu. Stąd te kryteria doboru badań muszą zagwarantować wystarczająco duże rozproszenie cech. Stanowi ono bowiem podstawę analizy statystycznej. Badania o zbliżonych lub nawet identycznych cechach uniemożliwiają analizę wpływu cech na miarę mocy efektu.

Opisywanie cech badań może wskazywać na te aspekty tematyczne i metodologiczne, które znalazły się w centrum badań, a więc zostały dobrze i szczegółowo przestudiowane oraz na zaniedbane, które poddane zostały płytkiej analizie.

Z uwagi na doniosłość pierwszego celu, którym jest statystyczna analiza cech, Hunter proponuje test wstępny, będący czynnością poprzedzającą kategoryzowanie i opisywanie. Test ten służy analizie rozproszenia rozkładu miar mocy efektu. W przypadku gdy wynik testu wskazuje, iż przynajmniej 75% wariancji pochodzi z błędu standardowego i błędu pomiaru, kategoryzowanie staje się procedurą niepotrzebną. Gdy wynik testu jest mniejszy niż 75%, kategoryzowanie ma swoje uzasadnienie (Hunter, Schmidt i Jackson 1982, s. 140).

Rozważania teoretyków problemu nasuwają następujące pytanie: które cechy należy kategoryzować? Empirycznej odpowiedzi na tak sformułowane pytanie należy poszukiwać w dotychczasowych metaanalizach i w zbiorze badań dobranych do metaanalizy. Problem cech dotyczących tematyki zasadniczo różni się od problemu cech metodologicznych.

5.1. Opis cech tematycznych

Teoretyczny punkt wyjścia prac badawczych nad metaanalizą stanowi podjęcie decyzji, a w jej konsekwencji – selekcja cech do opisu i dalszych analiz. Tym płaszczyznom działań towarzyszą głównie poszukiwania racjonalne, oparte na formułowaniu szczegółowych problemów badawczych. Wyborowi podlegają wszystkie te cechy, dla których uzasadniona jest hipoteza o ich istotnym wpływie na zmienne zależne. Zmienne te stanowią podstawę wyliczeń miary mocy efektu. Ilustrację omawianego zagadnienia dotyczącego doboru cech stanowią przytoczone przykłady:

- typ szkoły i rodzaj szkoły;
- metody nauczania;
- stosowane środki dydaktyczne;
- rodzaj podręczników;
- poziom wykształcenia nauczycieli;
- charakterystyka prób badawczych itd.

Drugiej części odpowiedzi na postawione wcześniej pytanie należy upatrywać w badaniach dobranych do metaanalizy. Z jednej strony wskazują one na szereg cech mających omówiony wpływ w pojedynczych przypadkach, z drugiej strony wytyczają granice badanej rzeczywistości. Niektóre z cech wybrane na podstawie teoretycznego punktu wyjścia metaanalizy mogą okazać się niedostępne w raportach pojedynczych badań. Badania wybrane do metaanalizy mogą stanowić istotne źródło doboru cech z uwagi na ich precyzyjną charakterystykę w sporządzanych raportach.

Ogromnym problemem przy opisywaniu omawianych cech jest brak danych. Raporty z pojedynczych badań często nie zawierają potrzebnych danych, co staje się powodem niedostrzegania wszystkich czynników, wpływających na ostateczne rezultaty. Skutkiem tego jest konieczność wyłączenia podobnych badań z analizy rozproszenia miary mocy efektu. Wyłączenie to z kolei może uniemożliwić sformułowanie odpowiedzi na najistotniejsze problemy badawcze metaanalizy.

5.2. Opis cech metodologicznych

Charakter prac badawczych, dotyczących cech metodologicznych wyróżnia się od badań, w których wnikliwym poszukiwaniom poddawane są cechy tematyczne. Istota tych różnic polega przede wszystkim na większej ogólności cech metodologicznych. W niewielkim stopniu także, w przeciwnie-

Opis cech badań

stwie do cech tematycznych, zależą od badawczego problemu metaanalizy. Niektóre cechy metodologiczne pojawiają się w licznych metaanalizach i w przypadkach niezwykle odmiennych problemów badawczych. Jako przykład można wyodrębnić kilka takich cech:

- modele eksperymentalne;
- stosowane techniki zbierania danych empirycznych;
- badania prowadzone na próbach lub na całych populacjach;
- użyte wskaźniki itd.

Doświadczenia dotychczasowych metaanaliz stanowią podstawę doboru cech metodologicznych w procesie opisu i kategoryzowania. Godne uwagi są propozycje Glassa i jego współpracowników, dotyczące schematów i recept kategoryzowania cech metodologicznych (Glass, McGaw i Smith 1981). Ilustracją przykładu tej czynności badawczej może być tabela do badań eksperymentalnych, sporządzona przez badaczy Kulika i Kulik (1989, s. 262).

Tabela 4. Cechy kategoryzowane w modelu metodologicznym

| Cecha | Kategorie |
|--|---|
| Sposób doboru grup porównawczych (eksperymentalnych i kontrolnych) | Dobór losowy lub celowy |
| Kontrola czynności nauczyciela | Ten sam nauczyciel dla grupy kontrolnej i eksperymentalnej lub różni nauczyciele |
| Kontrola czynników rozwoju | Jednoczesne eksperymenty we wszystkich grupach lub eksperymenty w różnych semestrach i latach |
| Kontrola jakości narzędzi badawczych | Testy standaryzowane lub niestandaryzowane |
| Kontrola obiektywizmu narzędzi badawczych | Zadania zamknięte lub otwarte |
| Kontrola sytuacji eksperymentalnej | Eksperymenty analizujące kowariancje lub eksperymenty niezawierające kontroli statystycznej |

ródło: Kulik i Kulik 1989.

Autorzy dokonali w niej zestawienia na podstawie założeń teoretycznych i empirycznego przeglądu ówczesnych metaanaliz. Tabela ta nie będzie szczegółowo omawiana, ponieważ większość wyodrębnionych w niej kategorii zostanie poddana szerszej analizie w kolejnych podrozdziałach.

Istotne w omawianej problematyce wydaje się pytanie: jakich doświadczeń dotyczących doboru i opisu cech badań dostarcza dotychczasowa praktyka metaanalizy? Najwięcej odkryć naukowych w tej dziedzinie przy-

noszą doniesienia autorów Kulika i Kulik (1989). Zajmowali się oni problemem współzależności cech badań i ich rezultatów. Wyraża się on w następujących pytaniach: jaki wpływ był badany najczęściej, jaka była jego siła i w jaki sposób można go wyjaśnić? Autorzy dokonali kwalitatywnej metaanalizy metaanaliz. Do swojego badania wybrali metaanalizy o dostatecznie dobrym poziomie metodologicznym. Dobór ten miał zagwarantować adekwatną rzetelność i trafność dociekań dokonanej metaanalizy.

Autorzy w swojej analizie wykorzystali wpływ opisanych i kategoryzowanych cech badań na moc efektu, otrzymaną w tych badaniach. Tabela 5 wskazuje na część otrzymanych rezultatów (Kulik i Kulik 1989, s. 271).

Wpływ czynności doboru grup eksperymentalnych i kontrolnych (próby badawcze w eksperymencie) okazał się istotnym, mimo iż w każdej z czterech metaanaliz nie był on jednakowy. Wszystkie badania pojedyncze zostały rozdzielone na dwie grupy: te, w których zastosowano dobór losowy (randomizacja), i te, w których nie zastosowano doboru losowego. Badania, w których zastosowano dobór losowy, były w większości badaniami o większej mocy efektu niż badania przeprowadzone z pominięciem procedury doboru losowego. Oznacza to, że w metaanalizie konieczne jest uwzględnianie czynności doboru prób do badań. Cecha ta powinna zostać poddana analizie statystycznej, a jej rezultaty należy odnieść do wyników ostatecznej interpretacji. Wpływ ten nie jest zupełnie systematyczny. Jedynie dwie z omówionych metaanaliz potwierdziły tę prawidłowość. Trzecia z nich dowiodła nieznacznego przeciwnego wpływu: badania z doborem losowym wykazały mniejszą moc efektu niż badania przeprowadzone z pominięciem procedury doboru losowego. W czwartej z omawianych metaanaliz nie zaobserwowano żadnych różnic. Mimo wyniku, wskazującego w jednej z metaanaliz na nieznaczący wpływ przeciwny, należy stwierdzić, że dotychczasowa praktyka metaanalizy potwierdziła wpływ doboru prób na otrzymane rezultaty oraz potrzebę uwzględniania tego wpływu.

Przeprowadzone badania dowiodły, że wpływ nauczyciela w eksperymencie jest niewiele mocniejszy. W niektórych badaniach wpływ ten był kontrolowany (we wszystkich grupach występował ten sam nauczyciel). Inne badania nie zawierały takiej kontroli (we wszystkich grupach występowali różni nauczyciele). Otrzymane wyniki wskazują, że badania, w których występował ten sam nauczyciel posiadają mniejszą moc efektu niż badania, w których występowali różni nauczyciele. Potwierdziły to trzy przeprowadzone metaanalizy, natomiast kolejna nie wskazała na żadne różnice. Fakt wystąpienia żadnych różnic w metaanalizie, w której efekty eksperymentalne okazały się silne, należy zaliczyć do korzystnych. Czwarta z omawianych meta-

Opis cech badań

Tabela 5. Wpływ niektórych cech badań na moc efektu, otrzymaną w tych badaniach

| cechy badań | metaanaliza I | | metaanaliza II | | metaanaliza III | | metaanaliza IV | |
|---|---------------|------------|----------------|------------|-----------------|------------|----------------|------------|
| | liczba badań | moc efektu | liczba badań | moc efektu | liczba badań | moc efektu | liczba badań | moc efektu |
| losowy dobór grup porównawczych | | | | | | | | |
| tak | 27 | 0,26 | 62 | 0,33 | 28 | 0,52 | 24 | 0,09 |
| nie | 67 | 0,32 | 86 | 0,29 | 73 | 0,52 | 91 | 0,21 |
| kontrola czynności nauczyciela | | | | | | | | |
| tak | 34 | 0,23 | 90 | 0,26 | 45 | 0,52 | 24 | 0,08 |
| nie | 57 | 0,35 | 49 | 0,42 | 53 | 0,52 | 91 | 0,21 |
| sposób publikowania badania | | | | | | | | |
| niepublikowane | 28 | 0,34 | 27 | 0,19 | 21 | 0,62 | 10 | 0,37 |
| dysertacje | 49 | 0,23 | 62 | 0,22 | 15 | 0,39 | 48 | 0,12 |
| badania publikowane | 17 | 0,46 | 59 | 0,44 | 65 | 0,52 | 57 | 0,20 |
| rodzaj narzędzi badawczych | | | | | | | | |
| testy nauczycielskie | 24 | 0,26 | 115 | 0,30 | 81 | 0,56 | 6 | -0,13 |
| testy nauczycielskie i ogólnie dostępne | 12 | 0,28 | 9 | 0,33 | 10 | 0,42 | 4 | 0,04 |
| testy ogólnie dostępne | 58 | 0,33 | 24 | 0,33 | 10 | 0,30 | 105 | 0,21 |
| rok wydania | | | | | | | | |
| przed 1960 | 0 | - | 0 | - | 0 | - | 29 | 0,26 |
| 1960–1969 | 9 | 0,36 | 10 | 0,12 | 1 | 0,81 | 67 | 0,20 |
| 1970–1979 | 53 | 0,32 | 99 | 0,31 | 78 | 0,49 | 17 | -0,03 |
| 1980 i później | 32 | 0,25 | 39 | 0,34 | 22 | 0,62 | 2 | 0,19 |

ródło: Kulik i Kulik 1989.

analiz objęła swym zasięgiem badania, których efekty eksperymentalne były niewielkie. W sytuacjach, w których średnie efekty eksperymentalne są niewielkie, różnica między badaniami zawierającymi kontrolę wpływu nauczyciela i badaniami pozbawionymi tej kontroli najczęściej jest skutkiem przypadku. Dlatego te jako potwierdzenie kontroli wpływu nauczyciela należy uznać dwie pierwsze metaanalizy. Kontrola wpływu nauczyciela zdecydowanie jest cechą, którą należy kategoryzować i kontrolować. Wpływ ten potwierdziły przeprowadzone metaanalizy badań eksperymentalnych. Rezulta-

ty badań wyłącznie eksperymentalnych nie stanowią często przedmiotu integracji w obszarze pedagogiki. Efekty nauczania porównuje się tak e w badaniach nieeksperymentalnych. W wielu z nich sytuacja dotycząca kontroli nauczyciela jest bardzo podobna jak w badaniach eksperymentalnych. Odnosi się to do takich badań, jak np. doskonalenie programu nauczania poszczególnych przedmiotów szkolnych, porównywanie efektywności rozmaitych metod nauczania i podręczników. Dlatego te dla rozwoju badań empirycznych w obszarze pedagogiki doświadczenia dotyczące wpływu nauczyciela są bardzo korzystne.

Sposób publikowania rezultatów badań wskazuje na jeszcze większe różnice. Nie występują one w sposób jednoznaczny. Trzy metaanalizy dowodzą, że publikowane badania (w książkach, w czasopismach itp.) wykazują większą moc efektu niż badania, które nie zostały opublikowane. W metodologii meta-analitycznej pytanie dotyczące wyjaśnienia tego zjawiska jest mniej ogólne. Każda kolejna metaanaliza musi poszukiwać rozwiązań przy uwzględnieniu natury zjawisk analizowanych w poszczególnych badaniach. Bardziej ogólne staje się pytanie poszukujące wyjaśnienia istnienia tego zjawiska. Jeżeli wpływ ten występuje, uzasadnione jest opisywanie i kategoryzowanie cechy oraz jej statystyczna analiza. W celu wyjaśnienia potrzeby uwzględniania sposobu publikowania wyników badań, należy powrócić do podejścia Slavina i Glassa. Według Slavina cecha ta może służyć jako kryterium doboru badań w metaanalizie. W pierwszej kolejności należy wybrać badania, którym można najbardziej zaufać. Jeżeli badania niepublikowane przeceniają moc efektu nie należy ich włączać do metaanalizy. Zdaniem Glassa (i większości autorów teoretycznych prac o metaanalizie, a także praktyków metaanalizy) cechę tę należy kategoryzować i analizować statystycznie. Przy założeniu, że wpływ publikowania jest niewielki, korzyść płynąca z przeprowadzonej analizy statystycznej jest mało istotna. Badania zawierające rezultaty bardziej rzetelne mają większe znaczenie i większy wkład w wyjaśnianie studiowanych zjawisk. Główną zaletą omawianego podejścia polega na tym, że z metaanalizy nie wyłącza się żadnego badania i żadnego rezultatu.

Z całościowego badania autorów (Kulik i Kulik) do interpretacji wybrano najistotniejsze wyniki. Do tabeli włączono wpływ dwóch czynników, w których nie dostrzeżono wprowadzić systematycznych efektów, ale stały się one istotne dla praktyki pedagogicznej.

Pierwszy z wymienionych to wpływ dotyczący techniki zbierania danych. Omówione cztery metaanalizy zawierały badania dydaktyczne, w których jako pomiar zmiennych zależnych zastosowano testy osiągnięć szkolnych uczniów. Przeprowadzone badania zostały rozwarstwione na trzy kategorie:

Opis cech badań

badania z testami ułożonymi przez samych nauczycieli, badania z testami ogólnie dostępnymi oraz badania, w których jednocześnie zastosowano testy nauczycielskie i ogólnie dostępne. Wyniki wskazały na istotne różnice. W niektórych przypadkach większą moc efektu wykazały badania z testami ogólnie dostępnymi, w innych natomiast większą moc efektu wykazały badania z testami sporządzonymi przez nauczycieli. Uwzględnianie cechy dotyczącej techniki zbierania danych w niektórych obszarach ma ogromne znaczenie, w innych zasadniczo mniejsze lub w ogóle żadnego, ponieważ niektóre różnice w tabeli są bardzo duże. Uboczny wniosek tej analizy posiada znaczenie w praktyce. Testy układane przez nauczycieli nie okazały się gorsze od testów powszechnie dostępnych. Jest to wynik specyfiki anglosaksońskiego obszaru językowego. W Stanach Zjednoczonych na rynku osiągalne są rozmaite testy. W większości są one dobrze wystandaryzowane w odróżnieniu od testów nauczycielskich, które wyjątkowo rzadko należą do standaryzowanych. Oznacza to, że badania nie potwierdziły powszechnie panującej opinii o niskim poziomie testów nauczycielskich. Należy jednak nadmienić, że w badaniach zawartych w omówionych czterech metaanalizach nie użyto testów sporządzonych dla potrzeb codziennej praktyki. Były one ułożone dla potrzeb badania eksperymentalnego. Z całą pewnością wpłynęło to na jakość sporządzanych testów. Potwierdza to kwalifikacje nauczycieli w dziedzinie sporządzania testów osiągnięć szkolnych uczniów. Wniosek ten ma duże znaczenie dla wykonawców badań pierwotnych.

Okres, w którym pojedyncze badania stosowane w metaanalizie były publikowane (wiek badań), nie miał wpływu systematycznego. Otrzymane różnice są wprawdzie większe niż w przypadku pozostałych cech. W pierwszej metaanalizie większą moc efektu wykazywały badania starsze, w drugiej przeciwnie – młodsze. Prawidłowość ta w dwóch prezentowanych metaanalizach jest klarowna i przejrzysta. Wynika to z wartości mocy efektu, która narasta wraz z wiekiem badania (w pierwszej metaanalizie) lub obniża się (w drugiej metaanalizie). W pozostałych dwóch metaanalizach sytuacja jest zupełnie przeciwna. Moc efektu nie zmienia się w sposób równomierny i zależy od wieku. Moc efektu najpierw obniża się, a potem narasta. Dlatego nie można jednoznacznie stwierdzić, że wiek publikacji zwiększa lub obniża moc efektu. Potwierdzenie wyraźnej prawidłowości mogłoby stanowić jedynie przegląd dużej liczby metaanaliz. Oznacza to, że nie potwierdziła się upowszechniona opinia zakładająca, że starsze badania są mniej rzetelne od badań młodszych. W obliczu otrzymanych wyników powstaje następujące pytanie: jak postępować w badaniach metaanalitycznych – kategoryzować badania, uwzględniając wiek publikacji, czy też pominąć procedurę katego-

ryzowania? Pierwsze empiryczne rezultaty nie dostarczają jednoznacznej odpowiedzi na to pytanie. Możliwe jest, że uwzględnianie wieku publikacji nie może wpłynąć w zasadniczy sposób na wnioski metaanalizy.

Kulik i Kulik rezultaty te interpretują trochę inaczej. Sądzą oni, że różnice te nie potwierdzają tezy o znaczeniu kategoryzowania omówionych cech. Twierdzą np. że brak zaufania do rezultatów *quasi*-eksperymentów (w porównaniu z prawdziwymi eksperymentami) nie jest uzasadniony. Pogląd tego nie można jednak respektować i akceptować w całej rozciągłości. Rezultaty wskazują przecie wyraźnie, że metaanaliza bez uwzględniania wpływu omówionych cech pomijałaby ważne aspekty studiowanych badań i została by bez licznych odpowiedzi istotnych dla praktyki i teorii.

Z drugiej strony niewielkie różnice wskazują, że znaczenie poszczególnych cech metodologicznych jest w dużej mierze uzależnione od tematyki badań. Należy zakładać, że określona cecha w różnych metaanalizach może mieć odmienne znaczenie.

Uwzględnianie cech badań nie jest główną czynnością metaanalizy. Jej rezultaty są konieczne do interpretacji efektów otrzymanych z poszczególnych badań. W wyniku tego są one pomocne w poszukiwaniu odpowiedzi na zasadnicze pytania metaanalizy. Rezultaty otrzymane przez Kulika i Kulik potwierdzają potrzebę opisu i kategoryzowania omówionych cech badań. Cechy objęte metaanalizami Kulika i Kulik, stanowią wstęp w procesie poszukiwania najbardziej istotnych własności. Autorzy opierali się na ówczesnej praktyce, wskazując sposób rozwiązywania tego problemu przez dokonane już metaanalizy. Oczekuje się, że przyszłe analizy praktyki metaanalitycznej wskażą ostatecznie na właściwości, które są konieczne w procedurze kategoryzowania. Problem ten z pewnością stanie się jednym z najbardziej kontrowersyjnych i najczęściej dyskutowanych dylematów teorii metaanalitycznej.

Sposobu interpretacji wpływu cech wymienionych i innych dotyczy ostatnie pytanie wypływające z tej analizy. Analiza statystyczna dostarcza jedynie podstaw wyjaśniających ten wpływ. Identyfikacja wpływu jest łatwiejszą częścią omawianego zadania. Dużo trudniej jest wyjaśnić sedno i wewnętrzne mechanizmy wpływu. Metaanaliza integruje rezultaty ogromnych ilości badań. Potwierdzają one często różne, a nawet przeciwne wnioski. Stosunkowo prosta w interpretacji jest sytuacja, w której wszystkie badania wykazują jednakowy kierunek wpływu. Nie ma znaczenia, czy wskazują one jednocześnie na podobną moc efektu. Zdecydowanie bardziej skomplikowana jest sytuacja, w której dane badania potwierdzają jeden kierunek wpływu, inne

Opis cech badań

badania – inny kierunek. W tych okolicznościach powstaje pytanie: jak często będą występowały pierwsze sytuacje, a jak często drugie?

Kłopoty te nie mogą jednak stanowić argumentu dla przyjęcia metodologii, która w ogóle nie koduje cech i nie analizuje ich wpływu.

Autorzy Kulik i Kulik szczegółowo analizowali dalsze cechy badań. Wyjaśnili oni wpływ rodzaju publikacji na rezultaty badań. Porównywali badania publikowane w czasopismach z dysertacjami, wychodząc z założenia, że rezultaty publikowane w czasopismach częściej potwierdzają założone hipotezy badawcze. Fakt ten można zinterpretować następująco: autorzy dysertacji są bardziej rygorystyczni i surowi przy weryfikacji swoich hipotez. Kulik i Kulik wskazane różnice tłumaczą tendencją badaczy, autorów przeglądów i redaktorów czasopism pedagogicznych do prezentowania i publikowania przede wszystkim rezultatów, które najbardziej odróżniają się od założonych oczekiwań i których efekty są największe. Dlatego też proponuje się wykonawcom badań metaanalizycznych, aby opierali je częściej na dysertacjach niż na czasopismach pedagogicznych.

Opisywanie i kategoryzowanie cech metodologicznych nie stanowi tak poważnego problemu, jakim jest kategoryzowanie i opisywanie cech tematycznych. Większość autorów prowadzących metaanalizy podkreśla, że problem braku danych jest równie mniej doniosły niż w przypadku cech tematycznych (McGaw 1988, s. 682). Doświadczenia autora niniejszego opracowania nie potwierdzają tych poglądów. Metaanaliza badań przeprowadzonych w Słowenii w latach osiemdziesiątych dowiodła, że zastosowana w niej analiza tematyczna okazała się dużo łatwiejsza i prostsza od analizy metodologicznej. Przy doborze danych do analizy tematycznej w ogóle nie pojawił się problem braku danych w raportach. Z kolei w przypadku analizy metodologicznej stanowił on potężną barierę. Główną przyczyną tego stanu były niedokładne opisy modelu metodologii badań (Kožuh 2005). Zjawisko różnic w prezentowanych wynikach zasługuje i wymaga dodatkowych studiów nad badanym problemem.



Rozdział VI

Kategoryzacja badań

6.1. Trafność

Trafność kategoryzowania jest właściwością oznaczającą, że kategoryzowanie zawsze pozwoli badaczowi osiągnąć zamierzony cel, czyli że dane otrzymane dzięki kategoryzowaniu dobrze odzwierciedlają prawdziwe właściwości badań.

Przy trafności kategoryzowania istotne dla badacza jest to, czy dane, otrzymane dzięki procedurze odzwierciedlają prawdziwe właściwości i cechy poszczególnych badań. Trafność i rzetelność kategoryzowania zależą od zdefiniowania kategorii i od stopnia subiektywności ich zastosowania. Istotnym czynnikiem jest również cecha określająca pełność raportów badań wykorzystanych w metaanalizie (wszystkie potrzebne dane, zaprezentowane czytelnie i przejrzysto). Jednak nawet najbardziej precyzyjne i jednoznacznie określone kategorie nie mogą same w sobie zagwarantować wysokiej trafności i rzetelności, jeżeli raporty badań nie zawierają wszystkich potrzebnych danych.

Trafność niezwykle rzadko można zbadać empirycznie i określić ją kwantytatywnie w postaci wskaźników. Nie można również jednoznacznie stwierdzić, że dana procedura jest trafna, a innej brak trafności metodologicznej. Trafność jest właściwością, którą charakteryzuje się każda procedura kategoryzowania w większej lub mniejszej mierze. Istotnym jej elementem jest czynność walidacji. Walidacja kategoryzowania stanowi proces, którego nigdy nie można ostatecznie zakończyć – zawsze można coś dodać lub ulepszyć.

Kategoryzowanie cech badań ma dwie ważne właściwości. Jedną z nich jest określenie, czy kategorie rozwarstwiania odzwierciedlają zasadnicze cechy badań. Drugą z nich polega na całkowitej i wyczerpującej liście wyodrębnionych w badaniu kategorii. Całkowitość ta określa, czy lista kategorii obejmuje wszystkie ważne aspekty badanego zjawiska.

Trafność rozwarstwiania oraz kategoryzowania można ustalić jedynie drogą racjonalną. Dla każdej procedury rozwarstwiania i kategoryzowania należy wykonać racjonalny test trafności. Niezwykle uważać nie trzeba również przestudiować wszystkie kategorie w każdej pojedynczej procedurze kategoryzowania.

Do ustalania trafności kategoryzowania można posłużyć się metodą ekspertów (ocen ekspertów). Racjonalnego testu trafności powinno dokonać kilku ekspertów. W świetle dotychczasowych teorii i ustaleń są oni w stanie ocenić adekwatność kategorii. Oznacza to, że dokonuje się oceny jasności kryteriów, jednoznaczności kategorii oraz stopnia ich zasięgu. Przy poszukiwaniach tego rodzaju można opierać się na dotychczasowych doświadczeniach przeprowadzonych w badaniach metaanalitycznych. Postulowany racjonalny test trafności ma dwa podstawowe cele. Pierwszym z nich jest poprawa trafności kategoryzowania. Drugi cel to proponowanie następnym badaczom wystarczająco dobrej oceny trafności zastosowanej procedury. Dlatego ten omawiany test jest konieczny w dwóch różnych fazach metaanalizy. W fazie planowania badania oraz przygotowania do kategoryzowania spełnia on funkcję poprawy procedury i zwiększania jej trafności. Jest to ewaluacja formatywna toku postępowania. Znaczenie jej jest szczególnie duże dla badacza, który dokonuje metaanalizy. W fazie następującej po czynnościach rozwarstwiania należy ponownie zastosować racjonalny test z ocenami ekspertów. Jest to ewaluacja sumująca, która powinna dostarczyć użytkownikom informacji o trafności kategoryzowania.

Włączenie opisu dwóch omawianych testów jest konieczne w ostatecznym raporcie metaanalizy. Opis ewaluacji formatywnej ma znaczenie głównie dla badaczy, którzy zamierzają zajmować się analizą wtórną (np. przy późniejszych szerszych metaanalizach, obejmujących identyczną tematykę oraz przy innych metaanalizach o tematyce wystarczająco zbliżonej). Opis ewaluacji sumującej jest istotny głównie przy użyciu i aplikacji dociekań metaanalizy w praktyce wychowawczej.

6.2. Rzetelność

Rzetelność, będąca jedną z cech kategoryzowania, polega na tym, iż przy wielokrotnym kategoryzowaniu tych samych badań otrzymane wyniki są identyczne lub bardzo do siebie zbliżone. Rzetelność kategoryzowania, podobnie jak trafność, jest ściśle powiązana ze stopniem zdefiniowania kategorii i obiektywnością ich zastosowania. Istotnym czynnikiem jest również cecha określająca kompletność raportów badań wykorzystanych w metaanalizie.

Interesującą ideę, wskazującą na pewne rozstrzygnięcia w zakresie trafności i rzetelności prezentują Kulik i Kulik. Proponują oni, aby każdy raport metaanalityczny zawierał „opis wszystkich wanych cech badań objętych metaanalizą. Opis tych szczegółów w raporcie metaanalizy umożliwi innym badaczom stwierdzenie, czy badania w metaanalizie były rozwarstwione na odpowiednie kategorie. Pozwala to dalszym badaczom uzupełnić (zrewidować) metaanalizę z poprawionym i rozszerzonym kategoryzowaniem cech. Literatura metaanalityczna nie może stanowić prawdziwie skumulowanego materiału, dopóki niemal wszystkie raporty metaanaliz nie będą zawierały takich szczegółów” (Kulik i Kulik 1989, s. 262–263).

Sytuacja dotycząca rzetelności procesu kategoryzowania jest inna niż w przypadku trafności. W niej test racjonalny odgrywa nieporównywalnie mniejszą rolę. Nie oznacza to jednak, że nie ma on zastosowania w ustalaniu rzetelności. Może on pełnić funkcję ewaluacji formatywnej i ewaluacji sumującej. Pierwsza z wymienionych ewaluacji ma zasadnicze znaczenie przy ulepszaniu procedury kategoryzowania. Ewaluacja sumująca prowadzona drogą racjonalną nie spełnia żadnych istotnych funkcji. Stąd też przy ustalaniu rzetelności na uwagę zasługuje przede wszystkim podejście empiryczne.

Metoda powtórzeń (dwukrotnego kategoryzowania) jest główną metodą ustalania rzetelności procedury kategoryzowania. Już raz rozwarstwione badania powinny zostać poddane powtarzanej procedurze w sposób zupełnie niezależny. Procedura ta prowadzi do podwójnych rezultatów. Umogliwiają one obliczenie stabilności wyników, która jest miarą rzetelności kategoryzowania. Czynności te pozwalają na zastosowanie kilku metod statystycznych. Najprostszym zabiegiem jest przeliczenie tych badań, które zostały kategoryzowane w dwóch próbach w taki sam sposób. Należy dokonać w nich wyliczeń procentowych. Drugą czynnością statystyczną, której można użyć do opracowania rezultatów podwójnego kategoryzowania, polega na wyliczeniu wskaźników zależności (korelacji lub kontyngencji).

6.2.1. Wartości procentowe jako wskaźniki rzetelności

Już raz rozwarstwione badania zostały poddane powtarzanej procedurze rozwarstwiania i otrzymano podwójne rezultaty. Pierwszym zabiegiem jest przeliczenie tych badań, które zostały kategoryzowane w powtarzanej procedurze tak samo jak w pierwszej oraz wyliczenie wartości procentowych.

Większy procent jednakowo kategoryzowanych badań wskazuje na większą rzetelność kategoryzowania. Słabością czynności polegającej na

Kategoryzacja badań

obliczaniu procentów jest uwzględnianie jedynie występowania różnic lub ich braku. Procenty te nie wskazują jednak na stopień różnic w czynnościach kategoryzowania. Mimo wymienionych niedociągnięć należy uznać, że omawiana metoda jest prosta w użyciu i interpretacji oraz nie wymaga od badacza skomplikowanych przygotowań. Posiada jeszcze jedną zaletę, którą stanowi uniwersalność, oznaczająca, że metodę obliczania procentów można zastosować wszędzie bez względu na rodzaj skal pomiarowych w tych zmiennych.

Rezultaty zaprezentowane w tabeli 6 dotyczą dwóch kodowań 208 badań według sposobu doboru próby. Wyrońcono w nim dwie kategorie:

- 1) dobór losowy;
- 2) dobór inny niż losowy.

Tabela 6. Rezultaty dwukrotnego kategoryzowania 208 badań według sposobu doboru próby

| Rezultat kategoryzowania | Liczba badań | % |
|----------------------------------|--------------|-------|
| Jednakowo kategoryzowane badania | 186 | 89,4 |
| Różnic kategoryzowane badania | 22 | 10,6 |
| Razem | 208 | 100,0 |

ródło: badania własne.

Wartość procentowa badań, będących jednakowo rozwarstwionymi przy dwukrotnym kategoryzowaniu wynosi 89,4%, natomiast wartość procentowa badań, które w drugim kategoryzowaniu zostały rozwarstwione inaczej niż w kategoryzowaniu pierwszym, wynosi 10,6%. Wysoki procent jednakowo rozwarstwionych badań wskazuje na wysoką rzetelność. Jednak w procedurze poprawy rzetelności kategoryzowania należy nadal poszukiwać przyczyn powodujących wynik 10,6%. Przy analizie tak nieskomplikowanej cechy, jaką stanowi sposób doboru próby badawczej, zakłada się, że wartość procentowa inaczej rozwarstwionych badań będzie mniejsza lub zbliżała się do wartości zero. Racjonalna analiza kodowań może wskazywać na to, czy poszukiwana przyczyna wyników 10,6% tkwi głównie w mało czytelnych opisach kategorii, czy też w obiektywnych okolicznościach raportów badań. Jeśli przyjęte przez metaanalizę kryteria i opisy kategorii są nieprecyzyjne, stwarza to bezpośrednią możliwość naprawy rzetelności kategoryzowania. W przypadku gdy różnice pochodzą głównie z nieczytelnych i niekompletnych raportów badań, oznacza to, że zdecydowana poprawa rzetelności nie jest możliwa. Stwierdzenie to jest jednak korzystne. W interpretacji rezultatów metaanalizy, dotyczących kontroli statystycznej

Kategoryzacja badań

omawianej cechy (sposób doboru próby), uwzględnianie stwierdzeń o rzetelności kategoryzowania i przyczynach błędów jest nie tylko dopuszczalne, ale nawet konieczne.

Nie należy zakładać, że przy kategoryzowaniu badań według cech metodologicznych, w każdym przypadku wystąpią jedynie dwie kategorie. Kiedy badania przemieszczają się tylko w jednym kierunku (w przypadku występowania dwóch kategorii), istnieje możliwość odrębnego przeliczenia badań jednakowo kategoryzowanych i wyliczenia wartości procentowej. W przypadkach gdy zostaną zastosowane więcej niż dwie kategorie, problem przeliczenia empirycznego wskaźnika rzetelności okazuje się bardziej skomplikowany. Jeśli cecha ma naturę zmiennej nominalnej, zaobserwowane różnice nie będą duże. Zasadniczym niebezpieczeństwem takich przypadków jest niewłaściwa interpretacja rezultatów.

Dokonywanie odrębnych przeliczeń dla pierwszego i drugiego kategoryzowania jest niewystarczające, ponieważ nie istnieje możliwość obliczenia wartości procentowej jednakowo kategoryzowanych badań. Stwarza to konieczność określenia rodzaju kategoryzowania dla każdego badania w odrębny sposób. Jedynie takie podejście pozwala otrzymać dane, potrzebne do obliczenia wartości procentowej jednakowo kategoryzowanych badań.

W tabeli 7 zaprezentowano rezultaty dwóch kodowań 233 badań według sposobu gromadzenia danych, w których wyodrębniono kategorie:

- 1) ankieta;
- 2) wywiad z kafełką odpowiedzi;
- 3) wywiad bez kafełki odpowiedzi.

Tabela 7. Rezultaty dwukrotnego kategoryzowania 233 badań według sposobu gromadzenia danych

| Rezultat kategoryzowania | Liczba badań | % |
|----------------------------------|---------------------|----------|
| Jednakowo kategoryzowane badania | 204 | 87,6 |
| Różnie kategoryzowane badania | 29 | 12,4 |
| Razem | 233 | 100,0 |

ródło: badania własne.

Wartość procentowa jednakowo kategoryzowanych badań w pierwszym i drugim kategoryzowaniu wynosi 87,6%. Dokonując głębszej analizy indywidualnych przypadków niejednakowych kodowań, można jednocześnie odnaleźć przyczyny tego zjawiska. Wnioski płynące z tej analizy służą jako ewaluacja formatywna lub sumująca.

Kategoryzacja badań

W przypadkach gdy cecha kategoryzowania jest zmienną porządkową, istota problemu niejednakowego kategoryzowania ulega zasadniczej zmianie. Dotyczy to rozmiarów różnic, które w tej sytuacji mogą zaistnieć. Dlatego ta procedura analizy statystycznej jest o wiele bardziej skomplikowana. Również interpretacja wskaźnika rzetelności nie jest łatwa. Możliwości rozstrzygnięcia tej kwestii jest wiele. Jako przykład zostaną zaprezentowane dwie z nich.

Pierwszą czynnością uwzględnia fakt polegający na tym, że różnice między wielkościami w skali porządkowej (ordynalnej) nie są równe. Oznacza to, że np. dwa, trzy... stopnie na jednym końcu skali nie są równe dwóm, trzem... stopniom na innym końcu skali. Przykład ten stanowi dowód, że nie należy przeliczać ilości stopni w różnicach między pierwszym i drugim kategoryzowaniem poszczególnych badań. Z drugiej strony informacja określająca, czy dane badanie było w obydwu przypadkach kategoryzowane jednakowo lub różnie, jest niewystarczająca. Zmienna porządkowa stwarza o wiele więcej możliwości. Dlatego trzeba zestawiać trzy kategorie. W pierwszej kategorii należy umieścić badania, które w pierwszym i drugim kategoryzowaniu zostały rozwarstwione jednakowo. Druga kategoria powinna zawierać badania, które przy drugim kategoryzowaniu zostały rozwarstwione do sąsiednich kategorii (różnica między pierwszym i drugim kategoryzowaniem wynosi jeden stopień). Trzecią kategorię powinny wypełnić badania, które podczas drugiego kategoryzowania zostały rozwarstwione do „niasąsiednich” kategorii (różnica między pierwszym i drugim kategoryzowaniem wynosi dwa stopnie lub więcej). Na podstawie tych czynności można wyliczyć wartości procentowe dla wszystkich trzech kategorii. Należy pamiętać, że wysoką rzetelność można uzyskać w przypadku, gdy wartość procentowa pierwszej kategorii jest możliwie największa, a trzeciej kategorii – najmniejsza.

Tabela 8. Rezultaty dwukrotnego kategoryzowania 188 badań według sposobu instruowania eksperymentatorów

| Rezultat kategoryzowania | Liczba badań | % |
|---------------------------------|--------------|-------|
| Brak różnic | 159 | 84,6 |
| Różnica jednego stopnia | 21 | 11,2 |
| Różnica dwóch lub więcej stopni | 8 | 4,2 |
| Razem | 188 | 100,0 |

ródło: badania własne.

W obydwu przypadkach kategoryzowania 84,6% badań zostało rozwarstwionych jednakowo. Pozostałych 15,4% badań rozwarstwiono inaczej.

Kategoryzacja badań

W pierwszej kolejności należy odnaleźć przyczyny 4,2% badań, które w drugim kategoryzowaniu zostały rozwarstwione zdecydowanie inaczej niż w kategoryzowaniu pierwszym. Kolejną czynnością powinna być analiza badań rozwarstwionych do sąsiednich kategorii. Analiza różnic (zwłaszcza tych większych) umożliwia zwiększenie rzetelności kategoryzowania, które nie dotyczy jedynie dokonywanej metaanalizy. Gdy przyczyny badanych różnic mają charakter bardziej ogólny i pochodzą z obiektywnych warunków kategoryzowania, analiza odmiennych kodowań może dostarczyć przesłanek mających szersze zastosowanie.

W przypadkach gdy występuje duży odstęp od kategorii kategoryzowania i gdy można zakładać, że stopnie między tymi kategoriami są w pewnej mierze zbliżone, istnieje możliwość przeliczenia stopnia różnicy między kodowaniami. Dla każdego badania należy przeliczyć, ile stopni wynosi różnica między pierwszym i drugim kategoryzowaniem. Powstaną w ten sposób następujące kategorie:

- 1) brak różnic;
- 2) różnica jednego stopnia;
- 3) różnica dwóch stopni;
- 4) różnica trzech stopni itd.

Na podstawie wyodrębnionych kategorii można przeliczyć wartości procentowe dla każdej z kategorii. Należy pamiętać, że wysoka rzetelność można uzyskać w przypadku, gdy wartość procentowa pierwszej kategorii jest możliwie największa, a pozostałych kategorii – najmniejsza.

Tabela 9. Rezultaty dwukrotnego kategoryzowania 211 badań według sposobu instruowania obserwatorów

| Rezultat kategoryzowania | Liczba badań | % |
|--------------------------|--------------|-------|
| Brak różnic | 179 | 84,8 |
| Różnica jednego stopnia | 17 | 8,1 |
| Różnica dwóch stopni | 10 | 4,7 |
| Różnica trzech stopni | 5 | 2,4 |
| Razem | 211 | 100,0 |

Źródło: badania własne.

Ze względu na występowanie w podobnych tabelach kilku wartości procentowych, ich interpretacja i dokonywanie porównań z pozostałymi kodowaniami w obrębie tej samej metaanalizy lub z innymi metaanalizami jest bardziej skomplikowane. W prezentowanej tabeli ułatwia ją jedynie fakt polegający na tym, że liczba badań, w których występuje większa różnica między

pierwszym i drugim kategoryzowaniem, ulega równomiernemu pomniejszeniu. Zastosowanie omawianej metody z wartościami procentowymi wszystkich występujących różnic jest prostsze, gdy celem badacza jest poprawa procedury kategoryzowania (ewaluacja formatywna). Zdecydowanie trudniejszą czynnością jest dokonywanie porównania kodowań kilku lub większej liczby cech. Opieranie się jedynie na pierwszej kategorii, mimo faktu, że jest najwłaściwsza – to czynność niewystarczająca. Może bowiem zdarzyć się sytuacja, w której w procedurze kategoryzowania określonej cechy, procent jednakowo kategoryzowanych badań jest większy (efekt pożądanym), lecz pozostała część inaczej kategoryzowanych badań jest skoncentrowana wokół większych różnic występujących w niższych kategoriach (efekt niepożądany). W kategoryzowaniu innej cechy procent jednakowo kategoryzowanych badań może być mniejszy, ale pozostałe badania są skoncentrowane wokół mniejszych różnic. Wraz ze wzrostem liczebności kategorii wzrastają też trudności w interpretacji wyników kategoryzowania. Bardziej odpowiednie w takich przypadkach jest zastosowanie współczynników korelacji lub kontyngencji.

W sytuacjach, gdy występuje większa ilość wartości procentowych, ich interpretacja jest bardziej złożona. Trudniej również dokonać porównań między tymi kodowaniami a wszystkimi pozostałymi w obrębie tej samej metaanalizy oraz w obszarze innych metaanaliz. Stąd też w opisanych przypadkach należy zastosować współczynniki korelacji lub kontyngencji.

6.2.2. Współczynniki zbieżności i korelacji jako wskaźniki rzetelności

6.2.2.1. Współczynnik zbieżności Pearsona

Współczynnik ten odnajduje zastosowanie jako wskaźnik rzetelności w przypadkach, gdy liczba kategorii jest większa, a przyjęta skala zalicza się do porządkowych.

Zaprezentowana tabela 10 wskazuje na wyniki dwukrotnego kategoryzowania 272 badań. Cechą kategoryzowania jest zmienna porządkowa, wyodrębniająca cztery kategorie (I, II, III, IV).

Stopień zbieżności został określony za pomocą współczynnika Pearsona. Do jego przeliczenia stosuje się następujący wzór:

$$C_{\max} = \sqrt{\frac{\chi^2}{n + \chi^2}}$$

Kategoryzacja badań

Tabela 10. Wyniki dwukrotnego kategoryzowania 272 badań według reprezentatywności prób

| Pierwsze kategoryzowanie ↵ | I | II | III | IV | Σ |
|-----------------------------|----|----|-----|----|-----|
| Drugie kategoryzowanie ↵ | | | | | |
| I | 37 | 3 | 0 | 0 | 40 |
| II | 9 | 69 | 10 | 6 | 94 |
| III | 3 | 6 | 76 | 11 | 96 |
| IV | 0 | 0 | 2 | 40 | 42 |
| Σ | 49 | 78 | 88 | 57 | 272 |
| $\chi^2=481,18$ $C=0,80$ | | | | | |

ródło: badania własne.

Obliczony współczynnik zbite ności Pearsona posiada wartość 0,80. Stanowi to jednocześnie wskaźnik rzetelności kategoryzowania. Otrzymana wartość wskazuje na średnią rzetelność kategoryzowania. Interpretacja współczynnika jest jednak uzależniona także od konkretnych okoliczności. Dla niektórych cech badań wartość 0,80 może być interpretowana jako zadowalająca – w innych przypadkach jako niska i niewystarczająca. Współczynnik zbite ności Pearsona jest prostszy w interpretacji niż wartości procentowe. Rzetelniejsze i łatwiejsze są również porównania z innymi badaniami. Słabością współczynnika zbite ności Pearsona jest jednak jego zależność od rozmiarów tabeli, z której został obliczony. Jego największa wartość, nawet w przypadku współzależności funkcyjnej nie osiąga wyniku równego 1. W metaanalizie oznacza to, że nie istnieje możliwość otrzymania wartości $C=1$ nawet przy założeniu, że wszystkie badania zostały w obydwu czynnościach kategoryzowania rozwarstwione jednakowo. Właściwość ta zostanie zinterpretowana przykładem o charakterze hipotetycznym.

Przy ponownym kategoryzowaniu 95 badań na trzy kategorie wszystkie badania zostały skategoryzowane w taki sam sposób, jak w pierwszym kategoryzowaniu.

Z tabeli 11 mającej trzy rubryki poziome i trzy pionowe została wyliczona wartość współczynnika zbite ności Pearsona.

Powierzchniowa analiza tak dokonanego kategoryzowania ocenia rzetelność na poziomie średnim, lecz rzeczywisty wynik rzetelności kategoryzowania jest o wiele bardziej zadowalający – rzetelność ta jest stu procentowa. Trudności w interpretacji współczynnika zbite ności są jeszcze większe w przypadkach, w których dokonuje się porównania rzetelności dwóch

Kategoryzacja badań

Tabela 11. Wyniki hipotetycznego dwukrotnego kategoryzowania 95 badań

| Pierwsze kategoryzowanie ↘ | I | II | III | Σ |
|-----------------------------|----|----|-----|----|
| Drugie kategoryzowanie ↙ | | | | |
| I | 21 | 0 | 0 | 21 |
| II | 0 | 43 | 0 | 43 |
| III | 0 | 0 | 31 | 31 |
| Σ | 21 | 43 | 31 | 95 |
| $\chi^2=190,00$ $C=0,82$ | | | | |

ródło: badania własne.

rónych cech (z różną liczbą kategorii). Komparacje takie potrzebne są zwłaszcza wtedy, gdy w metaanalizie dokonuje się kategoryzowania podobnych (lub identycznych) cech, uwzględnianych w innej metaanalizie.

Dlatego te wartości C są porównywalne między sobą jedynie w przypadku, gdy pochodzą z tabel o jednakowych wymiarach. W sytuacjach, gdy dokonuje się porównania procedur z różną liczbą kategorii (a więc z tabelami różnej wielkości), należy zamiast współczynnika C zastosować skorygowany współczynnik C_{corr} .

Czynność korekcji należy do skomplikowanych, zwłaszcza w przypadkach, w których występują tabele z niejednakową liczbą rubryk poziomych i pionowych. Występowanie w tabelach przy procedurze podwójnego kategoryzowania zawsze jednakowej liczby rubryk poziomych i pionowych umożliwia w konsekwencji dokonywanie obliczeń skorygowanego współczynnika zbitejności Pearsona według następującego wzoru:

$$C_{corr} = \frac{C}{C_{max}}$$

oraz:

$$C_{max} = \sqrt{\frac{m-1}{m}}$$

m – liczba kategorii w tabeli (liczba rubryk poziomych lub pionowych).

Kategoryzacja badań

Tabela 12. Wyniki dwukrotnego kategoryzowania 185 badań według sposobu doboru prób

| Pierwsze kategoryzowanie ↘ | I | II | III | Σ |
|--|----|----|-----|-----|
| Drugie kategoryzowanie ↙ | | | | |
| I | 53 | 8 | 5 | 66 |
| II | 10 | 48 | 11 | 69 |
| III | 5 | 7 | 58 | 50 |
| Σ | 86 | 71 | 38 | 185 |
| $\chi^2=146,12$ $C=0,66$ $C_{corr} = \frac{0,66}{\sqrt{\frac{3-1}{3}}} = 0,81$ | | | | |

ródło: badania własne.

Nieskorygowana wartość współczynnika wskazuje na niską rzetelność kategoryzowania. Wartość skorygowana jest zdecydowanie wyższa, co w konsekwencji zwiększa ocenę rzetelności dokonanej na jej podstawie. Przykład ilustruje korzyści, które można uzyskać, stosując wartości skorygowane. W przypadku tabel o rozbudowanych rozmiarach korzyść ta jest odpowiednio mniejsza.

Liczba kategorii w tabelach jest w większości mała, jedynie w incydentalnych przypadkach przekracza wartość 10. Dlatego tego godna uwagi staje się propozycja przedstawiająca tabelę służącą do szybkiego obliczania maksymalnej wartości C (tabela 13 na następnej stronie).

Przykład przeliczenia współczynnika zbitej wartości Pearsona z tabeli 12 na skorygowany współczynnik kontyngencji C_{corr} :

$$C=0,66$$

$$C_{corr} = \frac{0,66}{0,816} = 0,81$$

Tabela 13. Maksymalne wartości współczynnika zbite ności Pearsona

| Liczba kategorii | C_{\max} |
|------------------|------------|
| 2 | 0,707 |
| 3 | 0,816 |
| 4 | 0,866 |
| 5 | 0,894 |
| 6 | 0,913 |
| 7 | 0,926 |
| 8 | 0,935 |
| 9 | 0,943 |
| 10 | 0,949 |
| 11 | 0,953 |
| 12 | 0,957 |
| 13 | 0,961 |
| 14 | 0,964 |
| 15 | 0,966 |
| ∞ | 1 |

ródło: badania własne.

6.2.2.2. Współczynnik zbite ności Cramera

Problemy związane z porównywaniem różnych tabel kategoryzowania za pomocą współczynnika zbite ności Pearsona rozwiązuje posługiwanie się w takich sytuacjach współczynnikiem zbite ności Cramera. Jego zaletą jest niezależność od rozmiarów tabel. Stąd możliwe staje się dokonywanie porównań bez względu na okoliczności. Należy jednak pamiętać, że współczynniki zbite ności Cramera nie są bezpośrednio porównywalne ze współczynnikami zbite ności Pearsona. Dlatego istnieje bezwzględna konieczność używania albo jednych współczynników, albo drugich. Współczynnik Cramera oblicza się według następującego wzoru:

$$V = \sqrt{\frac{\chi^2}{n(m-1)}}$$

Otrzymany współczynnik zbite ności Cramera jest niewysoki (tabela 14), wskazuje przy powierzchniowej analizie na niski poziom rzetelności. W celu dokonania porównania wyliczono z tej samej tabeli współczynnik zbite ności Pearsona ($C = 0,68$) i skorygowany współczynnik Pearsona ($C_{\text{corr}} = 0,83$).

Kategoryzacja badań

Tabela 14. Wyniki dwukrotnego kategoryzowania 225 badań według stopnia standaryzacji narzędzi badawczych

| Pierwsze kategoryzowanie ↵ | I | II | III | Σ |
|-----------------------------|----|----|-----|-----|
| Drugie kategoryzowanie ⇨ | | | | |
| I | 49 | 7 | 8 | 64 |
| II | 9 | 32 | 9 | 50 |
| III | 7 | 7 | 97 | 111 |
| Σ | 65 | 46 | 114 | 225 |
| $\chi^2=190,55$ $C=0,65$ | | | | |

ródło: badania własne.

Przeprowadzone porównanie wyraźnie wskazuje na to, że współczynnik Cramera nie podlega bezpośredniemu porównaniu z pozostałymi współczynnikami. Właściwość tę należy uwzględnić, dokonując interpretacji rzetelności. Główną zaletą współczynnika Cramera polega na dokonywaniu porównań tabel o różnych rozmiarach w sposób prosty i nieskomplikowany.

6.2.2.3. Współczynnik zbieżności ϕ

W przypadkach gdy kategoryzowanie zostało dokonane jedynie według dwóch kategorii, zamiast współczynnika Cramera można zastosować współczynnik ϕ . Jest to ten sam współczynnik, przystosowany jedynie do użycia w tabelach 2x2. Zaletą współczynnika ϕ jest proste i szybkie dokonywanie obliczeń, według wzoru:

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

Współczynnik otrzymany w tabeli 15 wskazuje na niski poziom rzetelności kategoryzowania omówionej cechy. Układ tabeli wyraźnie wskazuje na 39 badań, które podczas pierwszego i drugiego kategoryzowania zostały inaczej rozwarstwione (12,7%). Stanowi to dowód na istnienie potrzeby odnalezienia i wyeliminowania przyczyn błędów pojawiających się przy czynnościach kategoryzowania.

Kategoryzacja badań

Tabela 15. Wyniki dwukrotnego kategoryzowania 308 badań według rodzaju obserwacji

| Pierwsze kategoryzowanie ↵ | A | B | Σ |
|--------------------------------|-----|-----|-----|
| Drugie kategoryzowanie ↵ | | | |
| A | 92 | 15 | 107 |
| B | 24 | 177 | 201 |
| Σ | 116 | 192 | 308 |
| $\chi^2=163,05$ $\phi=0,73$ | | | |

ródło: badania własne.

6.2.2.4. Współczynnik korelacji Pearsona

Istnieją przypadki, w których podczas kategoryzowania dysponuje się zmiennymi przedziałowymi (interwałowymi). Zdarza się to jednak w bardzo nielicznych sytuacjach. Należy wówczas używać współczynnika korelacji Pearsona jako wskaźnika rzetelności. Ze wszystkich wskaźników współzależności, współczynnik korelacji Pearsona jest najlepszy – najlepiej odzwierciedla istotę współzależności. Jednak stosowanie tego współczynnika w metaanalizie należy do rzadkości.

Tabela 16. Wyniki dwukrotnego kategoryzowania 167 badań według czasu trwania czynności eksperymentalnych

| Pierwsze kategoryzowanie ↵ | 1 | 2 | 3 | 4 | 5 | Σ |
|----------------------------|----|----|----|----|----|-----|
| Drugie kategoryzowanie ↵ | | | | | | |
| 1 | 30 | 2 | 0 | 0 | 0 | 32 |
| 2 | 6 | 19 | 5 | 0 | 0 | 30 |
| 3 | 2 | 2 | 19 | 3 | 2 | 28 |
| 4 | 2 | 2 | 6 | 20 | 6 | 36 |
| 5 | 0 | 0 | 0 | 2 | 39 | 41 |
| Σ | 40 | 25 | 30 | 25 | 47 | 167 |
| $r=0,90$ | | | | | | |

ródło: badania własne.

Kategoryzacja badań

Współczynnik korelacji zaprezentowany w tabeli 16 wskazuje na wysoki poziom rzetelności omówionej cechy. Należy jednak podkreślić, że w rzeczywistości zmienna powinna mieć charakter numeryczny (przynajmniej zbliżony do skali przedziałowej). Wyniki tabeli dowodzą, że 76% badań dwukrotnie kategoryzowano w taki sam sposób. Stanowi to stosunkowo niski procent. Wysoka wartość współczynnika korelacji jest wynikiem rozwarstwienia w sąsiednie kategorie większości różnie kategoryzowanych badań. Przykład ten wskazuje na konieczność zachowania ostrości przy interpretacji współczynników korelacji, zwłaszcza w przypadku dokonywania komparacji z wartościami procentowymi.

6.2.2.5. Stosunek korelacji

Kiedy z powodu nielinearności korelacji między rezultatami dwóch kodowań zastosowanie współczynnika korelacji Pearsona jest niemożliwe, można go zastąpić stosunkiem korelacji. Wartości omawianego wskaźnika przy niewielkiej nielinearności są porównywalne z wartościami współczynnika korelacji Pearsona. Stosunek korelacji nie ma znaku plus lub minus, dlatego też konieczna jest szczególna ostrość przy interpretacji wyniku. Do niezwyklej rzadkości należą jednak sytuacje, w których przy opracowywaniu danych z dwóch kodowań tych samych badań można otrzymać tak odmienne rezultaty, których skutkiem mogłaby być korelacja negatywna.

Tabela 17. Wyniki dwukrotnego kategoryzowania 167 badań według czasu trwania obserwacji

| Pierwsze kategoryzowanie ↘ | 1 | 2 | 3 | 4 | 5 | Σ |
|----------------------------|----|----|----|----|----|-----|
| Drugie kategoryzowanie ↗ | | | | | | |
| 1 | 24 | 5 | 0 | 0 | 0 | 29 |
| 2 | 6 | 27 | 4 | 0 | 0 | 37 |
| 3 | 0 | 5 | 20 | 6 | 0 | 31 |
| 4 | 0 | 0 | 2 | 41 | 4 | 47 |
| 5 | 0 | 0 | 0 | 3 | 20 | 23 |
| Σ | 30 | 37 | 26 | 50 | 24 | 167 |
| $\eta=0,94$ | | | | | | |

ródło: badania własne.

Kategoryzacja badań

Interpretacja wartości stosunku korelacji jest zbliżona do interpretacji współczynników korelacji Pearsona. Otrzymana wartość jest bardzo wysoka, należy jednak uwzględnić fakt, że rezultaty tabeli wskazują na 21% badań kategoryzowanych w inny sposób.

6.2.2.6. Tetrachoryczny współczynnik korelacji

W przypadkach, w których kategoryzowania dokonuje się według tabeli 2x2 i występuje w niej skala porządkowa (oznacza to występowanie dwóch kategorii np. więcej – mniej, lepiej – gorzej itp.), zamiast współczynnika ϕ można zastosować współczynnik tetrachoryczny. Jego podstawową zaletą stanowi niezwykle precyzyjne odzwierciedlenie istoty współzależności rezultatów. Należy jednak pamiętać, że w omawianych przypadkach nadal istnieje możliwość zastosowania współczynnika ϕ , ale stanowi to zdecydowanie gorszą alternatywę.

Tabela 18. Wyniki dwukrotnego kategoryzowania 167 badań według adekwatności wykorzystanej literatury

| Pierwsze kategoryzowanie ↙ | Mniej adekwatna | Bardziej adekwatna | Σ |
|----------------------------|-----------------|--------------------|----------|
| Drugie kategoryzowanie ↘ | | | |
| Mniej adekwatna | 48 | 11 | 59 |
| Bardziej adekwatna | 10 | 98 | 108 |
| Σ | 58 | 109 | 167 |
| $r_T=0,91$ | | | |

ródło: badania własne.

Otrzymana wartość jest wysoka i wskazuje jednocześnie na wysoki poziom rzetelności omówionej cechy. Trudno jednak nie dostrzec, że aż 13% badań zostało w powtórny sposób kategoryzowanych inaczej rozwarstwionych niż podczas pierwszego kategoryzowania.

6.2.2.7. Określanie rzetelności metodą reprezentatywną z rezultatów prób

W przypadkach, kiedy metaanaliza obejmuje ogromną liczbę badań pojedynczych, można zastosować metodę reprezentatywną. Należy wówczas ze wszystkich kategoryzowanych badań do kolejnego kategoryzowania wybrać mniejszą część (próbę). Na tej próbie należy dokonać powtórnego kate-

Kategoryzacja badań

goryzowania na podstawie kategorii u ytych w pierwszym kategoryzowaniu. Wyniki rezultatów obydwu kodowań umo liwiają otrzymanie informacji, z których mo na obliczyć procent badań jednakowo skategoryzowanych lub współczynnik korelacji (w zale ności od natury kategorii). Uzyskane wartości procentowe lub współczynniki korelacji stanowią ocenę z prób (wartości estymatora), z których nale y wyliczyć przedział ufności i w kolejnej fazie poziom istotności.

6.2.2.7.1. Ocenianie procentu badań jednakowo kategoryzowanych z próby

Pierwsze (główne) kategoryzowanie zostało dokonane na wszystkich badaniach dobranych do metaanalizy, natomiast drugie kategoryzowanie tylko na próbie liczącej 143 badania. Liczba badań jednakowo kategoryzowanych wynosiła 116. Powstaje pytanie: ile wynosi wartość procentowa takich badań w populacji generalnej?

Tabela 19. Wyniki dwukrotnego kategoryzowania 143 badań w próbie

| | |
|---|-------------|
| Liczebność próby | 143 |
| Liczba badań jednakowo kategoryzowanych | 116 |
| % badań jednakowo kategoryzowanych | 81,12 |
| Błąd standardowy oceny procentowej | 3,28% |
| Przedział ufności dla 5% ryzyka | 74,68–87,56 |

ródło: badania własne.

Wartość procentowa jednakowo kategoryzowanych badań w populacji wszystkich badań waha się w granicach między 74,68% i 87,56%.

Górna granica odgrywa mniejszą rolę, o wiele istotniejsza jest granica dolna (74,68%). Jest to minimalna wartość mo liwa do otrzymania (z założonym ryzykiem 5%). Wskazuje ona wyraźnie na fakt, e omawiana procedura kategoryzowania posiada rzetelność **przynajmniej** na poziomie 74,68%.

6.2.2.7.2. Ocenianie współczynnika korelacji Pearsona z próby

Z procedury podwójnego kategoryzowania 75 badań w próbie, otrzymano współczynnik korelacji Pearsona $r=0,87$. Jaka zatem jest wartość tego współczynnika w całej populacji badań? Odpowiedzi na tak postawione pytanie dostarcza tabela 20.

Z powodu du ych wartości współczynników korelacji w próbie (oraz realnego założenia, e współczynnik korelacji w populacji generalnej będzie wysoki), a tak e z uwagi na małą liczebność próby, nale y zastosować procedurę Fishera.

Kategoryzacja badań

Tabela 20. Wyniki podwójnego kategoryzowania 54 badań w próbie

| | |
|---|------------|
| Liczebność próby | n=75 |
| Współczynnik korelacji Pearsona | r=0,87 |
| Współczynnik Fishera | zr=1,333 |
| Błąd standardowy oceny współczynnika Fishera | 0,118 |
| Przedział ufności (z 5% ryzyka) dla wartości współczynnika Fishera | 1,10–1,564 |
| Przedział ufności (z 5% ryzyka) dla wartości współczynnika Pearsona | 0,77–0,92 |

ródło: badania własne.

Zaprezentowany przykład pozwala na stwierdzenie, że kategoryzowanie całej populacji badań (z 5% ryzykiem) posiada rzetelność wyrażoną za pomocą współczynnika korelacji, wynoszącą przynajmniej 0,77.

Fazy wstępne metaanalizy, czyli przegląd literatury badań i dobór badań, są zdecydowanie bardziej skomplikowane niż w dotychczasowych przeglądach kwantytatywnych. Przynajmniej tak samo lub jeszcze bardziej złożone jest opisywanie i kategoryzowanie tych cech badań, które zostaną w metaanalizie kontrolowane statystycznie. To właśnie w zasadniczy sposób odróżnia metaanalizę od zwykłych przeglądów. Aby stało się możliwe wyliczenie wagi rezultatów poszczególnych badań i pomiar ich wkładu w ostateczne ustalenia metaanalizy, należy ich podstawowe cechy opisać kwantytatywnie oraz analizować statystycznie. Dotyczy to zarówno metodologicznych, jak i tematycznych cech badań. Procedury kategoryzowania i opisywania dostarczają badaczowi potrzebnego materiału empirycznego. Kontrola statystyczna i analiza cech badań stanowią konieczny element metaanalizy. Wymaga to realizacji dwóch zasadniczych warunków. Do pierwszego z nich zalicza się identyfikację i opis cech badań. Warunek drugi stanowi wystarczająco duże rozproszenie tych cech.

W raporcie badawczym metaanalizy procedura kategoryzowania i opisywania powinna być zaprezentowana w taki sposób, aby jego odbiorca mógł dokonać oceny zakresu oraz kryteriów doboru, a także wskazać na możliwości uzupełnienia i rozszerzenia dokonanych czynności badawczych.

Rozdział VII

Moc efektu w badaniach

Istota badań metaanalitycznych polega na integracji rezultatów zbioru badań, które dotyczą tego samego problemu badawczego. Błędne jest jednak dokonywanie integracji jedynie z zastosowaniem uproszczonych procedur polegających na „sumowaniu” i „odejmowaniu”. Do kwantytatywnej integracji tych rezultatów konieczne są nowe metody statystyczne. Stanowią one narzędzie, dzięki któremu możliwe jest odnalezienie odpowiedzi na dwa zasadnicze pytania metaanalizy. Pierwsze z nich dotyczy określenia mocy efektu otrzymanej z pojedynczych badań. Pytanie drugie obejmuje problem wpływu poszczególnych cech badań na ich rezultaty. Kwerenda i analiza dotychczasowej praktyki badań metaanalitycznych wyraźnie wskazują, że wcześniejsze metaanalizy w większości nie określały mocy efektu i nie uwzględniały wpływu cech badań. W ostatnich latach uwzględnianie tych przesłanek stało się wymogiem standardowym. Dlatego też większość dokonywanych obecnie metaanaliz spełnia te wymogi. Określanie mocy efektu zwiększa rzetelność i trafność wniosków w metaanalizie. Dotyczy to również uwzględniania wpływu cech badań na ich rezultaty (np. liczebność prób, model eksperymentalny, sposób publikowania itd.). Problem określenia mocy efektu większość autorów postrzega jako konieczność w metaanalizie (McGaw 1988, Rosenthal 1991, Hedges 1984 itd.). Autorzy nie są jednak zgodni w ujmowaniu problemu, dotyczącego kontroli statystycznej oraz uwzględnianiu cech badań. Odmienność poglądów i opinii w tych kwestiach jest głównie skutkiem wniosków płynących z analiz empirycznych dokonanych metaanaliz. Wskazują one wyraźnie, że wpływ cech badań na ich rezultaty jest w większości przypadków niewielki (Glass, McGaw i Smith 1981).

Już od początku dwudziestego stulecia badań empirycznych dotyczących zjawisk wychowawczych dokonuje się prawie wyłącznie na próbach. Powodów takiego postępowania jest wiele. Na uwagę zasługują zwłaszcza trzy z nich. Pierwszy to zbyt duże populacje badawcze w obszarze pedagogiki, które uniemożliwiają ich całkowite objęcie zasięgiem badań. Powód drugi to

ogromne rozproszenie geograficzne populacji badawczych w obszarze pedagogiki, a zwłaszcza w obrębie systemu oświaty (nawet w przypadkach, gdy dokonuje się badań zjawisk o małej liczebności). Ostatni, trzeci z kluczowych powodów dotyczy faktu, że populacje badawcze w obszarze pedagogiki są często złożone z kilku następujących po sobie generacji uczniów, studentów lub innych podmiotów procesu dydaktycznego.

Rezultaty badań empirycznych są najczęściej wyrażone w postaci przedziału ufności oraz testów statystycznych hipotez zerowych (Kozłowska 1999, s. 114). Wczesniejsze metaanalizy dowodzą, że badacze dokonywali przeliczeń rezultatów tych testów w zbiorze integrowanych badań.

Kolejną czynnością badaczy stało się rozwarstwianie tych rezultatów najczęściej na trzy kategorie. Pierwsza z nich zawiera rezultaty istotne statystycznie w jednym kierunku (rezultaty potwierdzające przyjętą hipotezę). Kategorię drugą stanowią rezultaty istotne statystycznie w przeciwnym kierunku, czyli rezultaty potwierdzające hipotezę przeciwną. Ostatnia, trzecia kategoria to rezultaty niebędące istotnymi, czyli takie, które nie potwierdzają żadnej z przyjętych hipotez.

Wartość procentowa poszczególnych kategorii stała się podstawą opracowania wniosków końcowych. Słabością prezentowanego podejścia jest fakt, że poziom istotności nie zależy jedynie od oddziaływań badanych zjawisk, lecz także od liczebności prób. Powoduje to w konsekwencji, że niewielkie oddziaływanie może posiadać wysoki poziom istotności w przypadkach, gdy badanie zostało wykonane na wielkich próbach i odwrotnie – wielkie oddziaływanie może nie osiągać założonego poziomu istotności z powodu ujęcia do badań małych prób.

Czynność przeliczania rezultatów można zastosować również przy badaniach na całych populacjach. W tym przypadku stosuje się proste przeliczanie rezultatów, które potwierdzają lub obalają przyjętą hipotezę. Są to sytuacje, w których na ogół nie występuje średnia kategoria (brak rezultatów, które nie potwierdzają żadnej z przyjętych hipotez). Wyniki całych populacji są zgromadzone zawsze po jednej ze stron. W celu zwiększenia rzetelności i uniknięcia błędnej interpretacji należy jednak skorzystać z możliwości zastosowania kategorii średniej, która wskazuje na rezultaty leżące blisko środka i nie potwierdza wyraźnie żadnej prawidłowości. Najważniejszą słabością tej metody to, że uwzględnia ona jedynie kierunek efektu poszczególnych badań, nie uwzględnia natomiast mocy efektu.

Późniejsze metaanalizy przyniosły rozwiązania polegające na rozważaniu wartości procentowej badanych zmian w kierunku określonej hipotezy. W przypadku wartości procentowej można uwzględniać nie tylko kierunek,

ale również moc efektu. Oznacza to, że niewystarczające jest jedynie ustalenie, jak duża liczba wyników wartości procentowej potwierdza założoną hipotezę, ale istotne jest również to, z jaką siłą określona wartość procentowa potwierdza założoną hipotezę. Ilustruje to następujący przykład: wyniki 55% i 87% potwierdzają ten sam kierunek efektu, lecz istnieje ogromna różnica między mocą efektu, na którą wyniki te wskazują.

Ułomność opisanych czynności polega na tym, że większość z nich nie dokonuje pomiaru mocy efektu. Jeśli nawet wyniki te są uzyskiwane, to otrzymane miary mocy efektu są między sobą nieporównywalne. Staje się to powodem potężnych ograniczeń w procesie interpretacji wyników oraz podczas stawiania wniosków. Z uwagi na wymienione słabości, podejścia te nie będą stanowiły przedmiotu kolejnych analiz.

Dotychczasowe rozważania wskazują na konieczność odnalezienia metody, umożliwiającej określanie mocy efektu w jednakowej skali (niezależnie od jednostek pomiaru w poszczególnych badaniach). Wymóg ten wydaje się nieskomplikowany jedynie w aspekcie końcowych ustaleń. Nie jest to jednak proste zadanie z punktu widzenia różnorodności rezultatów empirycznych, koniecznych do otrzymania miar mocy efektu. Różnorodność ta jest skutkiem stosowania rozmaitych procedur pomiaru oraz różnych metod statystycznych przy badaniu tego samego problemu badawczego. Dlatego też rozsądny wydaje się postulat rozwoju kilku metod, uwzględniających różnorodność rezultatów empirycznych badań pierwotnych. Zdarza się, że różnorodność ta nie pozwala na włączenie wszystkich rezultatów do analizy statystycznej.

W badaniach empirycznych z obszaru pedagogiki jako rezultaty centralne, najczęściej pojawiają się trzy parametry:

- średnie arytmetyczne;
- współczynniki korelacji;
- wartości procentowe.

Obok średnich arytmetycznych i współczynników korelacji wyróżnia się liczne miary rozproszenia (głównie wariancje i odchylenia standardowe). Miary rozproszenia nie występują jednak nigdy w roli rezultatów centralnych, mimo że stanowią podstawę niemal każdej analizy statystycznej. Istnieje więc potrzeba opisu czynności pomiaru mocy efektu dla omówionych trzech parametrów. W większości metaanaliza integruje badania, które zawierają wyniki wyrażone w postaci tego samego parametru. Zdecydowanie rzadziej integracji podlegają badania, których rezultaty określają się za pomocą dwóch lub wszystkich trzech omówionych parametrów.

7.1. Miara mocy efektu średnich arytmetycznych

W pierwszej części podrzdział ten wska e na sytuacje, w których ró no-rodność rezultatów w badaniach pierwotnych jest niewielka. Podobieństwo rezultatów jest największe w przypadkach, gdy badania pierwotne wykazują zbie ność zastosowanego modelu metodologicznego. Klasycznym przykładem takiej sytuacji jest metaanaliza badań eksperymentalnych.

Eksperyment dydaktyczny nale y do najlepiej rozwiniętych form ekspe-rymentu pedagogicznego. Otrzymuje się w nim głównie dane wyra one w po-staci zmiennych numerycznych, które są interwalne (przedziałowe) lub bardzo do nich zbli one (np. liczba błędów w dyktandzie, liczba godzin poświęco-nych nauce w ciągu ka dego dnia, wyniki testów osiągnięć szkolnych uczniów, wyniki testów psychologicznych itp.). Dane te stanowią podstawę obliczeń średnich arytmetycznych i wariancji. Podstawowym celem stosowania śred-nich arytmetycznych oraz wariancji jest dokonywanie porównania między ró nymi grupami (klasy szkolne, szkoły, starsi i młodszy uczniowie, dziew-częta i chłopcy, studenci dzienni i zaocznicy itp.).

Integracji rezultatów badań eksperymentalnych nie mo na dokonać z wy-korzystaniem jednej metody statystycznej. W badaniach eksperymentalnych, dotyczących zbli onego problemu (lub niemal identycznego) u ywa się ró -norodnych modeli eksperymentalnych. Wynikiem tego jest większa liczba podejść statystycznych w określaniu miary mocy efektu.

Ju Glass w swoich pierwszych pracach z dziedziny metaanalizy wskazy-wał na jeden ze sposobów integracji badań eksperymentalnych. Jako miarę mocy efektu przyjął ró nicę standaryzowaną między średnimi arytmetycz-nymi grup eksperymentalnych i kontrolnych. Do tej standaryzacji zastoso-wał miarę rozproszenia. Sposób określania miary mocy efektu ukazuje na-stępujący wzór (Glass 1976):

$$d = \frac{\bar{y}_e - \bar{y}_k}{s}$$

| | |
|-------------|---|
| \bar{y}_e | średnia arytmetyczna grupy eksperymentalnej |
| \bar{y}_k | średnia arytmetyczna grupy kontrolnej |
| s | odchylenie standardowe |

Miara ta mo e słu yć jako ocena parametru w populacji generalnej. Jego wartość w populacji generalnej mo na wyrazić następującym wzorem:

Moc efektu w badaniach

$$D = \frac{M_e - M_k}{\sigma}$$

Co zawiera wskazana miara i w jaki sposób umożliwia dokonywanie porównań i integracji efektów z poszczególnych badań? Jest ona przede wszystkim oparta na skutkach. Jej podstawę stanowi różnica między osiągnięciami grupy eksperymentalnej i kontrolnej. Średnia arytmetyczna grupy eksperymentalnej wyraża efekty uzyskane w tej grupie, podobnie jak średnia arytmetyczna grupy kontrolnej wskazuje na efekty w swojej grupie. Różnica średnich arytmetycznych wyraża specyficzny efekt czynnika eksperymentalnego. Zdarza się jednak, że w badaniach pierwotnych przy dokonywaniu pomiaru bardzo zbliżonych lub identycznych zjawisk (zmiennych), znajdują zastosowanie rozmaite narzędzia. Różnice są więc takież jednostki pomiaru. W badaniach eksperymentalnych z obszaru pedagogiki niezwykle często jako zmienna zależna pojawia się poziom wiedzy uczniów. Najczęściej mierzy się go testami osiągnięć szkolnych (Kozłowska 2002, s. 16). W większości przypadków badacze samodzielnie tworzą testy do swoich badań. Konsekwencją tego jest fakt, że praktycznie w każdym badaniu stosuje się inne testy. Rezultaty testów wyrażane w postaci punktowej nie są między sobą bezpośrednio porównywalne. Wartości punktowe jednych testów odróżniają się zasadniczo od wartości punktowych innych testów. Należy dodatkowo zaznaczyć, że nawet wartości punktowe jednego testu nie są między sobą porównywalne. Przewaga jednego punktu w teście nie zawsze wskazuje na jednakowy wzrost poziomu wiedzy. Rezultaty testu nie stanowią skali przedziałowej. Dlatego te punkty z dwóch różnych testów nie są między sobą porównywalne. Przy komparacji wyników z dwóch testów, dodatkowo pojawiają się dwa inne problemy. Pierwszy z nich polega na braku w testach zera absolutnego. Problem drugi dotyczy różnorodnych stopni trudności poszczególnych pytań oraz trudności testu w ujęciu globalnym. Jest to przyczyną tego, że różnica między średnią arytmetyczną grupy eksperymentalnej i kontrolnej (bez uwzględnienia stopnia rozproszenia) nie może stanowić rzetelnej miary mocy efektu. Porównywanie i integrację otrzymanych efektów umożliwia odnalezienie sposobu dokonywania standaryzacji tej różnicy. Istnieje potrzeba wykonania działania polegającego na dzieleniu różnicy średnich arytmetycznych przez miarę rozproszenia. Miarę rozproszenia wyrażają te same jednostki pomiaru, które stanowią wynik średnich arytmetycznych. Dlatego ta miara mocy efektu jest niezależna od jednostek pomiaru. Ostatecznym rezultatem takiego podejścia jest standaryzowana miara mocy efektu.

Głębszego omówienia wymaga problem zastosowania miary rozproszenia, potrzebnej do standaryzacji różnicy średnich arytmetycznych. Istnieją trzy możliwe rozwiązania:

- 1) zastosowanie odchylenia standardowego w grupie eksperymentalnej;
- 2) zastosowanie odchylenia standardowego w grupie kontrolnej;
- 3) zastosowanie średniej wartości odchylenia standardowego obydwu grup (wewnątrz grupowego odchylenia standardowego, ang.: *pooled within group standard deviation*).

Odchylenie standardowe w grupie eksperymentalnej jest zdecydowanie najgorszym wariantem. W grupie tej rozproszenie uzależnione jest bezpośrednio od oddziaływań eksperymentalnych. Może ono być mniejsze lub większe (pod wpływem czynnika eksperymentalnego) w porównaniu z rozproszeniem występującym w grupie kontrolnej. Jeżeli oddziaływanie czynnika eksperymentalnego zwiększa rozproszenie, miara mocy efektu ulega zmniejszeniu i przeciwnie – jeżeli oddziaływanie czynnika eksperymentalnego zmniejsza rozproszenie, miara mocy efektu powiększa się. Skutkiem niedoceny mocy efektu jest zwiększona ostrość przy wyprowadzaniu wniosków końcowych. Stanowi to korzystny rezultat z punktu widzenia ryzyka generalizowania. Jednocześnie powoduje to często ukończenie badań pozbawionych generalizacji. Oznacza to, że słabość zastosowanej metody uniemożliwia formułowanie wniosków końcowych. Konieczna jest taka metoda, która w każdym przypadku, gdy istnieją obiektywne przesłanki wnioskowania, umożliwia drogę dojścia do nich.

Skutkiem przeceniania miary mocy efektu jest natomiast zwiększone ryzyko przy dokonywaniu generalizacji (ryzyko wysze, niż założone przez badacza). Zwiększa to możliwość pojawienia się błędnych wnioskowań, czyli takiego generalizowania, które nie jest oparte na rzeczywistej podstawie.

Zupełnie incydentalnie pojawiają się wystarczająco silne argumenty popierające hipotezę, zakładającą, że oddziaływanie czynnika eksperymentalnego nie zwiększa lub nie zmniejsza rozproszenia w porównaniu z grupą kontrolną. Oznacza to, że zastosowanie odchylenia standardowego w grupie eksperymentalnej jest możliwe tylko w wyjątkowych przypadkach.

Drugą możliwością jest zastosowanie odchylenia standardowego w grupie kontrolnej. Takie podejście proponował i stosował Glass. Jest ono argumentowane faktem, że rozproszenie w grupie kontrolnej nie zależy od oddziaływania czynnika eksperymentalnego. Jednocześnie więc nie jest ono uzależnione od zabiegu badacza. Stwierdzenie to dowodzi, że rozproszenie w grupie kontrolnej jest tym, które istniałoby bez względu na fakt prowadzenia eksperymentu. Dlatego też odchylenie standardowe grupy kontrolnej

właściwie spełnia funkcję standaryzowania miary mocy efektu. Większość badaczy w dotychczasowych metaanalizach stosowała rozproszenie grupy kontrolnej (Kulik i Kulik 1989).

Trzecim rozwiązaniem jest użycie średniej wa onej odchylenia standardowego obydwu grup. W licznych przypadkach rozwiązanie to stanowi jedyną możliwość (dzieje się tak w sytuacjach, gdy raporty badań pierwotnych nie zawierają wariacji grupy kontrolnej i eksperymentalnej). Praktyka dowodzi, że raporty niemal zawsze zawierają średnią wa oną wariację (wariację wewnątrzgrupową). Stosowanie tej wariacji jest uzasadnione także wymaga homogenicznością wariacji obydwu grup (McGaw 1988, s. 679–680). Kolejnym uzasadnieniem użycia średniej wa onej jest posiadanie przez tę wariację mniejszego błędu standardowego. Istnieje jednak argument przemawiający przeciwko jej stosowaniu. Wskazuje on, że omawiana wariacja zawiera częściowo wpływy czynnika eksperymentalnego.

Kwestia zastosowania odpowiedniej wariacji w procedurze standaryzacji różnic średnich arytmetycznych ciągle pozostaje dyskusyjna. Dostrzega się konieczność wnikliwej analizy jednego z omawianych argumentów, potwierdzających stosowanie średniej wa onej. Jest nim przesłanka dotycząca homogeniczności wariacji obydwu grup. Pojawia się ona także w niektórych metodach statystycznych stosowanych w badaniach, określanych mianem nieeksperymentalnych. Zaprezentowany przykład obejmuje analizę eksperymentu, uwzględniającego jeden czynnik eksperymentalny (zmienną niezależną), posiadający dwie kategorie. Omawiany eksperyment dokonywany jest przy podziale na dwie grupy: eksperymentalną i kontrolną. Miara mocy efektu postulowana przez Glassa w najprostszej postaci odnosi się do takiego właśnie eksperymentu (kolejne podrozdziały dowiodą, że miara ta znajduje zastosowanie także w innych eksperymentach). W takim eksperymencie kontrola wpływu wszystkich czynników związanych z poszczególnymi uczniami zostaje dokonana przez randomizację (dobór losowy). Oznacza to, że wybrana próba zostanie rozdzielona w sposób losowy na dwie grupy: eksperymentalną i kontrolną. Taki rozdział spowoduje, że grupy staną się bardzo wyrównane, choć niezupełnie identyczne. Między grupami pozostaną nietendencyjne różnice. Nie jest to jedyną możliwość kontroli wpływu omawianych czynników, ale jedna z najczęściej stosowanych. Po zakończeniu części praktycznej eksperymentu następuje procedura pomiaru zmiennej zależnej numerycznej. Przy opracowywaniu rezultatów takiego eksperymentu, pomocna jest najczęściej metoda statystyczna nazywana analizą wariacji. Zaprezentowane rozważania skłaniają do postawienia ogólnego pytania: czy efektywność dwóch kategorii czynnika eksperymentalnego jest odmienna?

W procedurze statystycznej pytanie to ulega konkretyzacji i brzmi: czy średnie arytmetyczne populacji generalnych (populacji reprezentowanych przez grupę kontrolną i eksperymentalną) są różne? Hipoteza zerowa zakłada, że średnie arytmetyczne są jednakowe (różnica równa się wartości zero). Jej weryfikacji dokonuje się za pomocą analizy wariancji.

W analizie wariancji należy określić dwie niezależne oceny rozproszenia. Pierwszą z nich wylicza się z rozproszenia wewnętrznego w grupach porównywalnych. Drugiej oceny dokonuje się poprzez wyliczenie rozproszenia między grupami porównywalnymi. Weryfikacja hipotezy zerowej bazuje na porównywaniu dwóch ocen rozproszenia (wariancji między grupami i wariancji w grupach). Jest to centralna procedura statystycznego opracowania danych w badaniach eksperymentalnych. Warunkiem realizacji omawianego podejścia jest identyczność rozproszenia rezultatów w populacjach generalnych. Brak spełnienia tego warunku uniemożliwia kontynuację opracowania statystycznego. Pociąga to za sobą konsekwencje polegające na braku możliwości dokonywania generalizacji na całych populacjach. Powoduje to w efekcie brak odpowiedzi na główne pytanie badawcze. W takim przypadku wszystkie rezultaty eksperymentu odnoszą się jedynie do grup, na których eksperyment został przeprowadzony.

Warunek identyczności wariancji populacji generalnych nie podlega bezpośredniej weryfikacji z uwagi na brak danych pochodzących z badanych populacji. Testowania dokonuje się za pomocą parametrów z prób. Słuzymy temu test statystyczny, który w analizie wariancji określamy jako test wstępny. Z wielu istniejących testów dwa stosowane są najczęściej: test Newman–Keulsa oraz test Cochra.

W przypadku gdy test kończy się obaleniem hipotezy zerowej, powstaje dowód, że omawiane wariancje różnią się między sobą. Oznacza to, że nie istnieje możliwość kontynuowania procedury analizy wariancji. W metaanalizie istotne jest to, że standaryzacja miary mocy efektu dokonana dzięki wariancji wewnątrzgrupowej w przypadku obalenia hipotezy zerowej jest nieuzasadniona.

Gdy przeprowadzany test kończy się brakiem obalenia hipotezy zerowej, oznacza to jednocześnie brak dowodu, że omawiane wariancje różnią się między sobą. Interpretuje się je wówczas jako identyczne, co oznacza, że dalsza procedura analizy wariancji jest uzasadniona. W takim przypadku do standaryzacji miary mocy efektu najlepiej zastosować wariancję wewnątrzgrupową. Jest ona lepsza od wariancji w grupie eksperymentalnej i w grupie kontrolnej.

W świetle dotychczasowych twierdzeń można konstatować, że wykorzystuje się głównie dwie miary mocy efektu. Pierwsza z nich jest owocem do-

Moc efektu w badaniach

ciekań naukowych Glassa, który zastosował odchylenie standardowe grupy kontrolnej (s_k):

$$d_{GLASS} = \frac{\bar{y}_e - \bar{y}_k}{s_k}$$

Druga miara mocy efektu polega na zastosowaniu wewnątrzgrupowego odchylenia standardowego (s_w). Miarę tę preferuje wielu metodologów (Hedges, Rosenthal, Wolf i inni). Wzór, według którego dokonuje się obliczeń, ma następującą postać:

$$d = \frac{\bar{y}_e - \bar{y}_k}{s_w}$$

Przypadki, w których są dostępne dane pochodzące z całych populacji generalnych, obligują badacza do uycia wariancji populacji kontrolnej. Dwie wcześniej omówione wariancje prób spełniają funkcję oceniającą wartości w badanych populacjach (konieczny jest wybór najlepszej oceny). Dowodzi to braku zasadniczych różnic, dotyczących sedna obydwu miar.

Dla praktyki nie pozostaje obojętne, która z miar podlega zastosowaniu. Okoliczności decydują o uyciu jednej z nich. Należy jednak unikać stosowania obydwu miar w jednej metaanalizie (choć taka możliwość nie jest wykluczona). Przedmiotem dalszych rozważań stanie się tylko druga z omawianych dotąd miar – „d”. Po uwzględnieniu niewielkich różnic strategia badawcza, która jej dotyczy, odnosi się także do miary proponowanej przez Glassa.

Istota prezentowanego podrozdziału polega na wskazaniu sposobów dokonywania obliczeń miary mocy efektu. Hedges dowodził, że omawiany sposób jej obliczania daje ocenę obciążoną, zwaną inaczej tendencyjną (Hedges 1981, Hedges i Olkin 1985). Dlatego też proponował metodę korygującą tę ocenę. Korekcja ta umożliwia otrzymanie nieobciążonej oceny miary mocy efektu według następującego wzoru:

$$d_u = \left(1 - \frac{3}{4(n_e + n_k - 2) - 1} \right) d$$

W celu szybkiego przeliczania wartości nieskorygowanych na skorygowane, używa się zwykle tabel. Dla wartości wyrażonej działaniem $(n_e + n_k - 2)$ należy odnaleźć w tabeli współczynnik J pomnożony przez nieskorygowaną miarę mocy efektu. Otrzymana w ten sposób wartość stanowi skorygo-

waną miarę mocy efektu. Tabela współczynników J znajduje się w aneksie (tabela D). Działanie to przeprowadza się według następującego wzoru:

$$d_u = d \cdot J$$

Niektórzy autorzy dowodzą, że wpływ korekcji dokonanej przez Hedgesa jest niewielki, a nawet zupełnie nieznaczny. McGaw uzasadnia, że w zwykłych okolicznościach różnica między oceną obciążoną i nieobciążoną jest nieznaczna (McGaw 1988, s. 680). Normalnymi okolicznościami określa on eksperymenty z liczebnością grup porównywalnych wynoszącą około dwudziestu osób. Takie Kulik i Kulik podkreślają mniejszy wpływ tej korekcji (Kulik i Kulik 1989). Autorzy analizowali empirycznie wpływ korekcji na wartość miary mocy efektu. Dokonali oni ponownej analizy statystycznej na niektórych swoich wcześniejszych metaanalizach. Dowiedli, że korelacja między korygowanymi i niekorygowanymi wartościami miary mocy efektu wynosi 0,99, a różnica między nimi w większości przypadków wynosi mniej niż 0,01 (Kulik i Kulik 1989, s. 244). Oznacza to, że wnioskowali oni na podstawie przeprowadzonych przez siebie badań, że twierdzenie Hedgesa, dotyczące procedury korekcji jest nieuzasadnione. Wnioskowanie to wskazuje jednak na mało krytyczną interpretację otrzymanych rezultatów.

Pomimo że korelacja między korygowanymi i niekorygowanymi wartościami miary mocy efektu może być bardzo wysoka, nie oznacza to wcale, że korekcja postulowana przez Hedgesa jest niepotrzebna (zwłaszcza w sytuacjach, gdy liczebność prób w badaniach objętych metaanalizą jest bardzo zbliżona). Przy niewielkich różnicach między liczebnością grup korelacja ciągle pozostaje wysoka. Wysoki wskaźnik korelacji nie oznacza, że korygowane wartości są identyczne (lub bardzo zbliżone) z wartościami niekorygowanymi. Takiej mocy argumentacji nie posiada aden współczynnik korelacji. Właściwość ta zostanie zaprezentowana na przykładzie krótkiej serii badań eksperymentalnych. Dla każdego badania obliczono korygowane i niekorygowane wartości. Otrzymane rezultaty pokazuje tabela nr 21.

Z rezultatów przedstawionych w tabeli został obliczony współczynnik korelacji Pearsona. Otrzymana wartość współczynnika wynosi 0,99. Mimo wielkich różnic między korygowanymi i niekorygowanymi wartościami miary mocy efektu otrzymany współczynnik korelacji jest wyjątkowo wysoki. Przykład ten nie dowodzi jednak, że różnice między korygowanymi i niekorygowanymi wartościami miary mocy efektu są zawsze tak ogromne. Celem przykładu jest bowiem relatywizacja dowodu o nieistnieniu potrzeby dokonywania korekcji sugerowanej przez Hedgesa. Oparcie krytyki Kulika i Kulika na wskaźnikach korelacji jest metodologicznie nieuzasadnione.

Tabela 21. Nieskorygowane i skorygowane wartości mocy efektu serii badań eksperymentalnych

| Badanie | Nie skorygowana wartość | Skorygowana wartość |
|---------|-------------------------|---------------------|
| 1 | 0,35 | 0,33 |
| 2 | 0,59 | 0,55 |
| 3 | 0,46 | 0,42 |
| 4 | 0,36 | 0,34 |
| 5 | 0,16 | 0,16 |
| 6 | 0,44 | 0,41 |
| 7 | 0,11 | 0,10 |
| 8 | 0,68 | 0,65 |
| 9 | 0,23 | 0,22 |
| 10 | 0,28 | 0,27 |
| 11 | 0,47 | 0,45 |
| 12 | 0,45 | 0,42 |
| 13 | 0,48 | 0,46 |
| 14 | 0,54 | 0,49 |
| 15 | 0,39 | 0,37 |
| 16 | 0,09 | 0,08 |
| 17 | 0,55 | 0,53 |
| 18 | 0,70 | 0,64 |

ródło: badania własne.

Zdecydowanie bardziej poważy jest argument drugi mówiący, że różnice między korygowanymi i niekorygowanymi wartościami miary mocy efekty w większości przypadków osiągają wartość mniej niż jedna setna. Dotyczy to jednak takich badań, w których liczebność grup porównywalnych wynosi około dwudziestu osób lub więcej. W badaniach eksperymentalnych próby mniejsze nie należą do rzadkości. Eksperymenty dydaktyczne prowadzone w jednej klasie szkolnej niemal zawsze obejmują próby mniejsze od dwudziestu. Są to eksperymenty, w których pojawiają się różnice większe niż wzmiankowane przez autorów Kulika i Kulik. Właśnie z powodu większych różnic w liczebności prób między poszczególnymi badaniami objętymi metaanalizą, Hedges rozwinął swoje podejście dotyczące przeprowadzania korekcji. Jest to kolejny kontrargument przeciwko krytyce autorów Kulika i Kulik. Chęć otrzymania przez badacza wniosków najbardziej rzetelnych, których nie obali lub nie podważy pierwsza przeprowadzona metaanaliza, wyzwała konieczność użycia korekcji Hedgesa.

Kolejny przykład (tabela 22) zawiera rezultaty metaanalizy dokonanej przez autorów. Wszystkie miary mocy efektu zostały skorygowane z założeniem

Moc efektu w badaniach

hipotetycznym, e zostały one otrzymane z różnych prób liczących mniej niż dziesięciu respondentów (przykład 1) oraz z prób, których liczba respondentów waha się między dwadzieścia i trzydzieści (przykład 2).

Tabela 22. Nieskorygowane i skorygowane wartości mocy efektu 24 badań

| Badanie | Nieskorygowana wartość | Skorygowana wartość 1 | Skorygowana wartość 2 |
|---------|------------------------|-----------------------|-----------------------|
| 1 | 0,52 | 0,48 | 0,51 |
| 2 | 0,33 | 0,31 | 0,33 |
| 3 | 0,48 | 0,44 | 0,47 |
| 4 | 0,71 | 0,66 | 0,70 |
| 5 | 0,44 | 0,41 | 0,43 |
| 6 | 0,66 | 0,61 | 0,65 |
| 7 | 0,47 | 0,44 | 0,46 |
| 8 | 0,55 | 0,51 | 0,54 |
| 9 | 0,35 | 0,33 | 0,35 |
| 10 | 0,45 | 0,42 | 0,44 |
| 11 | 0,70 | 0,62 | 0,68 |
| 12 | 0,29 | 0,27 | 0,29 |
| 13 | 0,62 | 0,59 | 0,61 |
| 14 | 0,28 | 0,26 | 0,28 |
| 15 | 0,49 | 0,46 | 0,48 |
| 16 | 0,81 | 0,75 | 0,80 |
| 17 | 0,44 | 0,42 | 0,43 |
| 18 | 0,36 | 0,33 | 0,35 |
| 19 | 0,21 | 0,19 | 0,21 |
| 20 | 0,49 | 0,46 | 0,48 |
| 21 | 0,64 | 0,59 | 0,63 |
| 22 | 0,56 | 0,52 | 0,55 |
| 23 | 0,40 | 0,38 | 0,39 |
| 24 | 0,38 | 0,36 | 0,37 |

ródło: badania własne.

Różnice między wartościami korygowanymi i niekorygowanymi występujące w mniejszych grupach są ogromne, sięgając w poszczególnych przypadkach różnicy dochodzącej nawet do 10%. W przypadku większych prób różnice są zasadniczo mniejsze i w niektórych przypadkach nie osiągają nawet jednej setnej. Przykład ten prowadzi do ustaleń ostatecznych, polegających na tym, że błędy spotykane przy obliczaniu miary mocy efektu nie są tak zupełnie nieznaczące. Stanowi to dowód na to, że korekcja proponowana

przez Hedgesa jest potrzebna i uzasadniona. Najbardziej korzystnym rozwiązaniem byłoby więc stosowanie korekcji Hedgesa we wszystkich możliwych przypadkach. Innym rozwiązaniem jest określenie granicy empirycznej do stosowania tej korekcji. W tym ujęciu należy pominąć dokonywanie korekcji przy większych próbach, a odnosić ją jedynie do prób mniejszych. Ustalenie punktu przebiegu proponowanej granicy jest bardzo trudne. Według dotychczasowych doświadczeń autorów Kulika i Kulik mogłaby ona przebiegać przy liczebności próby poniżej dwudziestu respondentów. Zgromadzone wcześniej argumenty skłaniają jednak do przesunięcia punktu granicznego do liczby około trzydziestu respondentów. Dylematy te rozwiązać może jedynie przyszła metaanalizyczna praktyka badawcza.

Kwestia przeprowadzania korekcji Hedgesa lub jej pomijania jest bardzo delikatna w obszarze nauk pedagogicznych. Eksperymenty pedagogiczne prowadzi się często na próbach o małej liczbie respondentów (przed wszystkim uczniów). Przeprowadzanie eksperymentów na większych próbach jest związane ze słabszą kontrolą sytuacji eksperymentalnej. Przykładem takich badań są eksperymenty przeprowadzane w kilku równoległych klasach na terenie tej samej szkoły lub w kilku klasach równoległych na terenie różnych szkół (każda szkoła reprezentowana przez jedną klasę). Bez względu na rozwój metodologii eksperymentu pedagogicznego w przyszłości eksperymenty przeprowadzane w jednej klasie będą ciągle potrzebne i aktualne. Dotyczy to również innych rodzajów badań pedagogicznych, w których występują małe grupy formalne lub nieformalne.

Interesujące, że metody określania mocy efektu z eksperymentów prowadzonych na dwóch porównywalnych grupach (eksperymentalnej i kontrolnej) są bardzo dobrze rozwinięte. Ciągłym niezbadanym obszarem pozostaje kwestia postępowania dotyczącego określania mocy efektu w przypadku badań, w których użyto bardziej kompleksowych modeli eksperymentalnych. Dotyczy to głównie modeli, w których dokonuje się porównania kilku kategorii czynnika eksperymentalnego. Modele te mają odpowiednio więcej grup poddanych komparacji. Zjawisko to występuje jeszcze bardziej wyraźnie w przypadku eksperymentów obejmujących swoim zasięgiem więcej niż jeden czynnik eksperymentalny. Eksperymenty prowadzone na podstawie dwóch czynników, posiadają przynajmniej cztery grupy porównywalne. Stosunkowo często zdarza się, że grup porównywalnych jest więcej niż cztery. Obecnie takie modele eksperymentalne używane są tak często, jak proste modele z dwiema grupami porównywalnymi.

Istnieje potrzeba odnalezienia właściwego rozwiązania tego problemu. Występowanie większej liczby grup porównywalnych powoduje jednocześnie

wystąpienie większej liczby średnich arytmetycznych i wariancji. Okoliczności te rodzą pytania:

- które średnie arytmetyczne należy poddawać procedurze porównywania?
 - czy dokonywać porównywania wszystkich średnich, czy tylko niektórych?
 - jeśli procedurze komparacji mają być poddane tylko niektóre średnie arytmetyczne, to w jaki sposób przeprowadzić ich selekcję?
 - które z wariancji należy zastosować w obliczeniach miary mocy efektu?
- Podobnych pytań powstaje jeszcze wiele, te jednak należą do najbardziej istotnych.

Odpowiedzi na postawione pytania należy poszukiwać w określeniu problemu badawczego metaanalizy. Ze szczególną stanowczością należy podkreślić, że nie dotyczy to problemów badawczych badań pierwotnych objętych metaanalizą. Dobór badań do metaanalizy powinien opierać się na precyzyjnie zdefiniowanym problemie badawczym. Problem ten sprawia wrażenie oczywistego, ale praktyka dowodzi jego złożoności. Stanowi on jednocześnie źródło pierwszych poważniejszych krytyk metody metaanalitycznej. Dotyczą one tego, że dokonane metaanalizy wiążą z sobą sprzeczne zjawiska. Różnorodność tematyczna badań objętych metaanalizą nie może zagrozić jednoznaczności i klarowności problemu badawczego metaanalizy.

Przypadki, w których faza ostatecznego zdefiniowania problemu badawczego metaanalizy jest dokonana ze szczególną troską, nie oznaczają jeszcze wcale, że zostały rozwiązane wszystkie problemy powiązane z różnorodnymi modelami badań eksperymentalnych. W poszczególnych badaniach tego typu pojawiają się różne czynniki eksperymentalne. Można odnaleźć w nich również te same czynniki, różniące się między sobą zastosowanymi kategoriami. Oznacza to, że dobór rezultatów rozmaitych badań należy ograniczyć zgodnie z pytaniami badawczymi metaanalizy. W konsekwencji tego postulatu do metaanalizy zostaną dobrane niektóre badania w całości, a niektóre tylko częściowo. Badania objęte metaanalizą mogą zawierać rozmaite kategorie czynnika eksperymentalnego. Należy jednak pamiętać, że w metaanalizie zawsze wybiera się tylko określone kategorie (lub bardzo zbliżone do określonych). Efektem takich założeń jest dobór jedynie tych rezultatów (np. różnic między średnimi arytmetycznymi), które powstały w grupach poddanych działaniom kategorii badanych w metaanalizie. Zakładając ograniczenie do np. trzech określonych kategorii, należy wybrać średnie arytmetyczne trzech wybranych kategorii z każdego badania. Pozostałe wyniki poszczególnych badań nie zostaną użyte w prowadzonej metaanalizie, ponieważ nie stanowią żadnych przesłanek do konstruowania odpowiedzi na badawcze pytania metaanalizy.

Ostateczną konsekwencją postulowanych stanowisk jest konieczność wyboru w każdym przypadku jednej pary grup porównywalnych (i ich średnich arytmetycznych) z badań prowadzonych także w szerszym zakresie (tzn. z badań zawierających kilka grup porównywalnych). Każda para wybrana z poszczególnych badań jest częścią składową zbioru par, dotyczących dwóch określonych kategorii. Wyniki otrzymane w tym zbiorze stanowią podstawę sformułowania odpowiedzi na określone pytanie badawcze. Liczba zbiorów poszczególnych par może być dowolna, lecz jednocześnie zgodna z obszarem problemu badawczego. W takich przypadkach miara mocy efektu zostanie obliczona według wzoru:

$$d = \frac{\bar{y}_1 - \bar{y}_2}{s}$$

gdzie:

\bar{y}_1 i \bar{y}_2 – średnie arytmetyczne wybranych grup porównywalnych,
 s – miara rozproszenia.

Dotychczas analizie poddano problem doboru miary rozproszenia w przypadkach prostych modeli eksperymentalnych. Na początku omawianego rozdziału wskazano na trzy istniejące możliwości jej selekcji. W modelach kompleksowych zaprezentowanych w tej części dostrzega się nowe możliwości. Stosowanie wariancji pierwszej lub drugiej z wybranych grup pozostaje nadal wykonalne. W eksperymentach obejmujących więcej niż dwie grupy zanika podział na grupy eksperymentalne i kontrolne. Wszystkie z nich są jednocześnie eksperymentalnymi i kontrolnymi. Z uwagi na to argumenty przemawiające za słusnością stosowania wariancji pierwszej lub drugiej grupy nie są tak jednoznaczne. Należy jednocześnie podkreślić, że proste ujęcie eksperymentu pozostawia możliwość określenia jednej z grup jako eksperymentalnej – drugiej jako kontrolnej. W takim przypadku możliwość staje się zastosowanie wariancji jednej z wybranych grup. W omawianych sytuacjach istnieją jednak lepsze rozwiązania. Do dyspozycji pozostają jeszcze dwie wariancje. Pierwszą z nich stanowi wariancja wewnątrzgrupowa, obliczona ze wszystkich użytych grup do eksperymentu. Druga wariancja jest także wewnątrzgrupowa, lecz obliczona tylko z pary wybranych grup.

Zaletą pierwszej wariancji jest występowanie mniejszego błędu standardowego, bez względu na jakiegokolwiek okoliczności. Jednak jej zastosowanie jest dyskusyjne, ponieważ uwzględnia ona grupy porównywalne, które być może nie będą w ogóle włączone do metaanalizy. Jeżeli krok po kroku zastosowano by rezultaty wszystkich grup porównywalnych, to użycie pierw-

szej wariacji byłoby uzasadnione. Stałaby się ona wówczas najlepszą oceną rozproszenia ogólnego (rozproszenia występującego w sytuacji pozbawionej dokonywania zabiegu eksperymentalnego). Jednocześnie byłaby najmniej obciążoną miarą rozproszenia. Jest to jednak mo liwość istniejąca tylko w teorii. Wymaga ona bowiem sytuacji, w której wszystkie badania objęte metaanalizą posiadają ten sam czynnik eksperymentalny z tymi samymi kategoriami. Prawdopodobieństwo zaistnienia takiego przypadku jest praktycznie zerowe. Opisywana wcześniej sytuacja, dotycząca postępującego u ycia wszystkich grup, mo e zdarzyć się jeszcze w jednym przypadku. Polega on na ograniczeniu metaanalizy jedynie do badań mających taki sam czynnik eksperymentalny z tymi samymi kategoriami. Ograniczenie to pozbawia badacza szeregu badań i rezultatów, które mogą odgrywać wielką rolę przy formułowaniu odpowiedzi na pytania badawcze metaanalizy. Zawę ona w ten sposób metaanaliza miałaby du o mniejszy poziom uogólnień i nowych stwierdzeń teoretycznych.

Wymienione przesłanki wskazują na korzyści płynące z zastosowania drugiej wariacji (wariacji wewnątrzgrupowej obliczonej jedynie z dwóch wybranych grup). U ywając jej, nale y jednocześnie zaakceptować większy błąd standardowy. Głównym problemem zastosowania tej wariacji jest brak mo liwości odnalezienia ka dorazowo potrzebnych danych w raportach poszczególnych badań. Wariację wewnątrzgrupową, odnoszącą się do wszystkich grup porównywalnych zawiera prawie ka dy raport badań eksperymentalnych w przeciwieństwie do wariacji poszczególnych grup porównywalnych, które w raportach występują rzadko. Raporty sporadycznie zawierają również dane surowe (nieopracowane). Dlatego te w metaanalizie bardzo często występuje sytuacja, w której nie ma mo liwości wyboru. Rozwinięta metodologia badań metaanalizy i szerokie zastosowanie tych badań w praktyce przyniosą najprawdopodobniej nowe rozwiązania w zakresie sporządzania raportów badań pierwotnych. Nale y oczekiwać, e w przyszłości raporty będą często zawierały również te rezultaty, które nie są bezpośrednio potrzebne do interpretacji, lecz korzystne dla późniejszych metaanaliz. Jednak oczekiwania te dotyczą przede wszystkim szerszych raportów w postaci monografii, elaboratów, dysertacji doktorskich itp. W czasopiśmie i pozostałych periodykach naukowych nie istnieje realna mo liwość takich zmian. Nierealne są również założenia oczekujące od redaktorów czasopiśmie naukowych, publikowania artykułów obciążonych mnogością rezultatów przygotowanych do ewentualnych późniejszych metaanaliz (metaanaliz, które być mo e nigdy nie zostaną przeprowadzone). Ka dy autor, który swoje doniesienia publikuje w czasopiśmie naukowych, nie tylko zna kryteria wymagań redak-

torów, ale wie również, jak trudno oczekiwać od nich opublikowania wszystkich potrzebnych do interpretacji rezultatów (z uwagi na ograniczenia w objętości czasopism). Dlatego gromna część raportów nie zawiera i nie będzie zawierała wszystkich danych, których oczekuje metaanalizy. Przypadki, w których założeniem metaanalizy jest zastosowanie określonej wariancji (a wszystkie raporty jej nie posiadają), prowadzą do ograniczenia metaanalizy jedynie do niektórych badań. Do metaanalizy należy włączyć wszystkie badania dotyczące określonej tematyki – tak jak doniesienia z badań prowadzonych wcześniej. Jednak te starsze raporty (nawet bardzo obszerne) nie dostarczają badaczom wszystkich potrzebnych do metaanalizy danych.

7.2. Miara mocy efektu współczynników korelacji

Integracja badań, w których jako główny wynik przyjęto współczynniki korelacji, jest zdecydowanie mniej skomplikowana od omówionych poprzednio. Mogą to być zarówno badania eksperymentalne, jak i nieeksperymentalne (nie jest to istotne w procedurze określania mocy efektu ze współczynników korelacji). Współczynnik korelacji jest niezależny od jednostek pomiaru. Dlatego nie wymaga on żadnej standaryzacji. Za miarę mocy efektu można przyjąć bezpośrednio współczynniki korelacji lub różne wskaźniki wyprowadzone z tych współczynników. Zależy to od konkretnych okoliczności dokonywanej metaanalizy.

7.2.1. Współczynnik korelacji Pearsona

Jedną z możliwości jest użycie współczynnika korelacji Pearsona jako miary mocy efektu. Wskaźniki te bardzo rzetelnie odzwierciedlają współzależność zmiennych numerycznych. Otrzymana w ten sposób miara mocy efektu jest dobrze porównywalna z wcześniej opisanymi miarami, przekształconymi na współczynniki korelacji. Podobnie jak w poprzednich tabelach druga potęga współczynnika korelacji (współczynnik determinacji) wskazuje na to, jak duża część wariancji zmiennej zależnej jest skutkiem wpływu zmiennej niezależnej. Przy dokonywaniu integracji badań eksperymentalnych zmienną niezależną jest czynnik eksperymentalny, natomiast w przypadku integracji badań nieeksperymentalnych zmienna zależna jest zupełnie dowolna (lecz zgodna z problemem badawczym metaanalizy). W przypadku stosowania współczynnika korelacji należy mieć na uwadze rozkład współczynników

korelacji otrzymanych na próbach. Współczynniki korelacji prób mają normalny rozkład jedynie w przypadku, gdy spełnione są następujące warunki (Nowaczyk 1985, s. 152–157):

- 1) próby powinny być wystarczająco duże;
- 2) wartość współczynnika korelacji w populacji generalnej powinna równać się wartości zero.

Praktyka uniemożliwia uwzględnianie warunku dotyczącego wartości zerowej współczynnika korelacji w populacji generalnej, ponieważ nie są znane dane o populacji generalnej. Można natomiast stosować współczynnik korelacji próby. W tym przypadku warunkiem umożliwiającym do realizacji są: duże próby, w których „ r ” mieści się w granicach między $\pm 0,80$, albo te małe próby, gdy „ r ” równa się zero (Nowaczyk 1985, s. 153). Przy stosowaniu badań z użyciem współczynnika korelacji próby zaleca się dużą ostrożność. Dlatego te warunki określa się bardziej rygorystycznie. Pierwszy z nich dotyczy liczebności próby badawczej, która powinna przekraczać stu respondentów. Drugim warunkiem jest występowanie współczynnika korelacji próby mniejszego od $\pm 0,50$. Przy spełnieniu dwóch wymienionych warunków zakłada się, że wartość współczynnika korelacji populacji generalnej jest bliska zero, a współczynniki korelacji prób posiadają normalny rozkład. Praktyka dowodzi, że spełnienie wskazanych warunków jest rzadkością, zwłaszcza wariant bardziej rygorystyczny. Stąd też bezpośrednie użycie współczynnika korelacji jako miary mocy efektu ma charakter incydentalny.

7.2.2. Współczynnik Fishera

W przypadku gdy „ r ” oddala się od zera i zbliża się do wartości 1,0, wówczas rozkład z próby oddala się od postaci normalnej i staje się coraz bardziej ujemnie skośny. Z tego względu opracowano statystykę Fishera. Umieżliwia ona transformowanie współczynnika „ r ” na współczynnik Fishera „ z_r ”. Współczynniki te posiadają normalny rozkład z próby nawet przy małych wielkościach prób. W przeciwieństwie od współczynnika korelacji, współczynnik Fishera waha się w granicach od $-\infty$ do $+\infty$. Przekształcenie wartości „ r ” na współczynnik Fishera można dokonać się według wzoru:

$$Z_r = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

lub:

$$Z_r = 1,1513 \log_{10} \left(\frac{1+r}{1-r} \right)$$

W praktyce częściej niż podanego wzoru używa się tabel do przekształcania wartości „r” na „Z_r” i odwrotnie – wartości „Z_r” na „r” (tabele E i F w aneksach).

W rzeczywistości badawczej najczęściej spotyka się sytuacje, w których wartości współczynników korelacji populacji generalnych są większe od zera, a próby badawcze są niewielkie. Zdarza się, że czasami badaniom poddaje się zalewności między zjawiskami, aby wykazać brak korelacji między nimi. Są to jednak niezwykle rzadkie przypadki. Konsekwencją tej rzadkości jest niemal całkowity ich brak w badaniach metaanalizacyjnych. Pytania i hipotezy badawcze w większości koncentrują się wokół poszukiwania i wyjaśniania współzależności. Pozwala to na przypuszczenie, że w praktyce metaanalizycznej podejście Fishera będzie stosowane bardzo często.

Badania pierwotne zawierają najczęściej współczynniki Pearsona i Fishera. W badaniach, w których występują jedynie współczynniki Pearsona, należy dokonać przekształcenia ich na współczynniki Fishera. Współczynniki Fishera służą wówczas jako miara mocy efektu.

7.2.3. Miara mocy efektu Rosenthala

Rosenthal definiuje jako jedną z możliwych miar mocy efektu stosunek między współczynnikiem korelacji Pearsona i współczynnikiem alienacji (pierwiastek $1-r^2$). Miara ta nie posiada żadnych większych zalet w porównaniu z bezpośrednim użyciem współczynnika korelacji Pearsona. Główną jej zaletą jest jednak to, że jest bardzo przydatna w testowaniu hipotez zerowych (test poziomu istotności). W celu dokonania testu Studenta konieczne jest jedynie działanie mnożenia jej przez liczbę stopni swobody ($n-2$).

Zastosowanie miary mocy efektu Rosenthala nie rozwiązuje jednak zasadniczo żadnego z omawianych wcześniej problemów, dotyczących bezpośredniego użycia współczynnika Pearsona. Tym samym nie dostrzega się większej potrzeby jej stosowania.

Obliczenia miary mocy efektu Rosenthala dokonywane są na podstawie wzoru:

$$ES = \frac{r}{\sqrt{1-r^2}}$$

7.2.4. Pozostałe współczynniki korelacji

Omawiane dotychczas możliwości stosowania współczynnika korelacji dotyczą sytuacji, gdzie metaanaliza integruje badania, w których konkretne warunki (charakter zmiennych i charakter ich współzależności) umożliwiają zastosowanie współczynnika korelacji Pearsona. W badaniach pedagogicznych częściej używa się jednak pozostałych wskaźników korelacji niż współczynnika Pearsona. Przyczyną tego jest fakt, że większość zjawisk wychowawczych należy mierzyć skalą porządkową. Jako przykład wymienia się kilka zjawisk tj.: motywacja, osiągnięcia, klimat, zainteresowania, satysfakcja, komunikacja, itp. Powoduje to, że w metaanalizie jako miary mocy efektu odnajdują zastosowanie pozostałe współczynniki korelacji: współczynnik Spearmana, stosunek korelacji, współczynnik korelacji dwuseryjny itd.

Przeomyślenia dotyczące zastosowania współczynnika Pearsona jako miary mocy efektu odnoszą się w dużej mierze również do pozostałych współczynników korelacji. Większość pozostałych współczynników może być dobrze porównywalna ze współczynnikiem Pearsona. Dlatego też wycie oraz interpretacja tych współczynników są zbliżone do zastosowania i omawiania współczynnika Pearsona.

7.2.5. Różnica współczynników korelacji

Zdarza się, że metaanalizą obejmuje się badania zawierające współczynniki korelacji, odnoszące się do dwóch lub więcej grup porównywalnych. Jeśli efekt danej czynności definiowany jest jako różnica między współczynnikami korelacji, wówczas jako miarę mocy tego efektu można przyjąć kilka różnych wskaźników. Wszystkie te wskaźniki pochodzą z omawianych różnic. Jako przykład zostaną zaprezentowane dwie główne miary mocy efektu:

1) różnica dwóch wskaźników korelacji:

$$ES = r_1 - r_2$$

2) różnica dwóch współczynników Fishera:

$$ES = Zr_1 - Zr_2$$

Z przedstawionymi wyżej miarami związane są wszystkie przeomyślenia zgromadzone wcześniej na temat zastosowania oraz interpretacji współczynników Pearsona i Fishera, a także pozostałych współczynników korelacji.

7.3. Miara mocy efektu wartości procentowych

W pedagogice częściej spotyka się jakości niż ilości, tym samym w badaniach częściej występują zmienne porządkowe oraz zmienne nominalne (skale ocen, wielostopniowe odpowiedzi w kwestionariuszach, rangi, stopnie wykształcenia, płeć, motywacja, itp.). W takich przypadkach nie istnieje możliwość stosowania średnich arytmetycznych, wariancji lub współczynników korelacji. Rezultaty badań w wymienionych sytuacjach wyrażone są najczęściej w postaci wartości procentowych. Przed dokonaniem obliczeń miary należy przekształcić wartości procentowe na proporcje. Wartości procentowe są łatwiejsze w interpretacji, lecz jednocześnie obciążone zbędnymi działaniami matematycznymi. Dlatego też w metaanalizach częściej używane są proporcje. Proporcje i wartości procentowe nie muszą podlegać procedurze standaryzacji, ponieważ należą one do wskaźników relatywnych. Ich użycie jest tak samo nieskomplikowane, jak stosowanie współczynników korelacji. Oczywiście miary mocy efektu otrzymane z wartości procentowych lub proporcji nie są bezpośrednio porównywalne z miarami uzyskanymi ze współczynników korelacji. Sytuacja ta w jeszcze większym stopniu dotyczy miar otrzymanych z różnic średnich arytmetycznych. Porównania te nie mają jednak większego znaczenia w dokonywaniu metaanaliz. Trudno sobie wyobrazić metaanalizę badań pierwotnych, w których w celu rozwiązania tych samych problemów badawczych jedni badacze używaliby danych ilościowych, drudzy stosowaliby dane jakościowe. Do pewnej miary założenie to jest możliwe. Wprawdzie niezwykle rzadko, ale może wystąpić konieczność integracji badań zawierających dane ilościowe i jakościowe. W takich przypadkach należy odnaleźć w badaniach pierwotnych dane surowe i dokonać przekształcenia np. zmiennych przedziałowych na porządkowe oraz stosować je wraz z pozostałymi zmiennymi porządkowymi. Podobna sytuacja dotyczy zmiennych nominalnych. Taka procedura pozbawia jednak badacza istotnych informacji, zawartych w danych oraz w rezultatach. Jej wielką zaletą jest to, że umożliwia dokonanie integracji wszystkich rezultatów. Alternatywą tej strategii badawczej jest dokonywanie odrębnych metaanaliz. Oznacza to tym samym wykluczenie metaanalizy wszystkich rezultatów. Omawiana sytuacja wymaga precyzyjnej oceny i dokonania „bilansu strat i zysków”, polegającego na wyborze utraty części informacji lub utraty możliwości dokonania całkowitej metaanalizy. Z pewnością w praktyce metaanalitycznej są to przypadki spotykane wyjątkowo rzadko. Badana rzeczywistość incydentalnie wymusza dokonywanie omówionych wyborów. Z uwagi na przedstawione argumenty,

przeprowadzanie komparacji różnic miar nie ma większego znaczenia dla metaanalizy.

W tej części zostaną zaprezentowane dwie możliwe miary mocy efektu wykorzystywane w celu integracji wartości procentowych.

1. W sytuacjach, gdy poszczególne badania pierwotne posiadają tylko jedną grupę badawczą (bez względu na to, czy jest to próba lub populacja generalna) jako miarę mocy efektu stosuje się różnicę między proporcją danego zjawiska a wartością 0,50, czyli:

$$ES = p - 0,50$$

Otrzymane w taki sposób miary mocy efektu wahają się w granicach między -0,50 (maksymalny negatywny efekt) i +0,50 (maksymalny pozytywny efekt). Wartość 0,00 oznacza, że nie zaobserwowano żadnego efektu. Przesłanką użycia tej metody jest założenie, że 0,50 (czyli 50%) stanowi zerowy punkt referencji. Punkt ten określa stan, który występowałby w badaniu, gdyby nie pojawił się żaden efekt badanego czynnika. Dlatego ten efekt mierzy się jako odległość otrzymanego rezultatu od tego punktu. Zdefiniowana w ten sposób miara mocy efektu może posiadać również negatywne wartości. Jeżeli w konkretnej metaanalizie istnieją argumenty przemawiające za doбором innego punktu wyjścia, należy wówczas wybrać ten punkt, a nie wartość 0,50. Punkt wyjścia 0,50 jest standardowy, ale nie stanowi żadnej konieczności.

2. W sytuacjach, gdy poszczególne badania pierwotne posiadają dwie porównywalne grupy badawcze (podobnie jak w eksperymentach) jako miarę mocy efektu stosuje się różnicę między proporcjami tych dwóch grup, czyli:

$$ES = p_1 - p_2$$

Otrzymane w taki sposób miary mocy efektu wahają się w granicach między -1,00 (maksymalny negatywny efekt) i +1,00 (maksymalny pozytywny efekt). Wartość 0,00 oznacza, że w przeprowadzonym badaniu nie zaobserwowano żadnego efektu. Z powodu dokonania w tej mierze komparacji dwóch grup nie dostrzega się potrzeby stosowania zerowego punktu referencji. Proporcja w jednej grupie jest punktem referencji dla proporcji w grupie drugiej – i odwrotnie. Podobne zjawisko omówiono już w części dotyczącej badań eksperymentalnych.

7.4. Interpretacja mocy efektu

W przypadku gdy rezultaty w grupie eksperymentalnej i kontrolnej kształtują się według normalnego rozkładu (lub zbliżonego do normalnego) powstaje możliwość ustalenia bezpośrednio z wartości miary mocy efektu stosunku między rezultatami z obydwu grup. Wartości procentowe otrzymuje się z tabeli krzywej normalnej (tabela C w aneksach).

Przykład 1.

Wartość $d = 0,00$ oznacza, że średni rezultat w grupie eksperymentalnej przekracza 50% rezultatów w grupie kontrolnej, czyli rezultaty w grupie eksperymentalnej są na ogół takie same jak w grupie kontrolnej.

Przykład 2.

Wartość $d = 0,29$ oznacza, że średni rezultat w grupie eksperymentalnej przekracza 61,41% rezultatów w grupie kontrolnej.

Przykład 3.

Wartość $d = -0,57$ oznacza, że średni rezultat w grupie eksperymentalnej przekracza 38,43% rezultatów w grupie kontrolnej, czyli rezultaty w grupie kontrolnej są wyższe niż w grupie eksperymentalnej.

Przykład 4.

Wartość $d = 0,11$ oznacza, że średni rezultat w grupie eksperymentalnej przekracza 54,38% rezultatów w grupie kontrolnej, czyli rezultaty w grupie eksperymentalnej są bardzo zbliżone do rezultatów w grupie kontrolnej.

Przykład 5.

Wartość $d = 0,70$ oznacza, że średni rezultat w grupie eksperymentalnej przekracza 75,80% rezultatów w grupie kontrolnej.

Przykład 6.

Wartość $d = 1,57$ oznacza, że średni rezultat w grupie eksperymentalnej przekracza 94,18% rezultatów w grupie kontrolnej, czyli prawie wszystkie rezultaty w grupie eksperymentalnej są wyższe niż rezultaty w grupie kontrolnej.

Przykład 7.

Wartość $d = 2,61$ oznacza, że średni rezultat w grupie eksperymentalnej przekracza 99,55% rezultatów w grupie kontrolnej, czyli praktycznie wszystkie rezultaty w grupie eksperymentalnej są wyższe niż rezultaty w grupie kontrolnej.

W celach łatwiejszego i szybszego ustalania wartości procentowych sporządzona została krótsza tabela, która prezentuje wartości procentowe odpowiadające określonym wartościom „d”.

Moc efektu w badaniach

Tabela 23. Wartości procentowe odpowiadające wartościom „d”

| Wartość „d” | % | Wartość „d” | % |
|-------------|-------|-------------|-------|
| 0,00 | 50,00 | -0,00 | 50,00 |
| 0,05 | 51,99 | -0,05 | 48,01 |
| 0,10 | 53,98 | -0,10 | 46,02 |
| 0,15 | 55,96 | -0,15 | 44,04 |
| 0,20 | 57,93 | -0,20 | 42,07 |
| 0,25 | 59,87 | -0,25 | 40,13 |
| 0,30 | 61,79 | -0,30 | 38,21 |
| 0,35 | 63,68 | -0,35 | 36,32 |
| 0,40 | 65,54 | -0,40 | 34,46 |
| 0,45 | 67,36 | -0,45 | 32,64 |
| 0,50 | 69,15 | -0,50 | 30,85 |
| 0,60 | 72,57 | -0,60 | 27,43 |
| 0,70 | 75,80 | -0,70 | 24,20 |
| 0,80 | 78,81 | -0,80 | 21,19 |
| 0,90 | 81,59 | -0,90 | 18,41 |
| 1,00 | 84,13 | -1,00 | 15,87 |
| 1,10 | 86,43 | -1,10 | 13,57 |
| 1,20 | 88,49 | -1,20 | 11,51 |
| 1,30 | 90,32 | -1,30 | 9,68 |
| 1,40 | 91,92 | -1,40 | 8,08 |
| 1,50 | 93,32 | -1,50 | 6,68 |
| 1,75 | 95,99 | -1,75 | 4,01 |
| 2,00 | 97,72 | -2,00 | 2,28 |
| 3,00 | 99,87 | -3,00 | 0,13 |

ródło: badania własne.

Wskazaną strategią można interpretować każdą wartość miary mocy efektu. Należy jednak uważać, która z wariacji została użyta do obliczenia miary mocy efektu. Przedstawiona interpretacja dotyczy przypadków, w których użyta została wariancja grupy kontrolnej. W przypadkach zastosowania drugiej wariacji należy przemyśleć precyzyjnie, na co wskazują otrzymane wartości procentowe. Wprawdzie nie wskazują one na coś zupełnie odmiennego, ale istnieje potrzeba dużej ostrożności przy formułowaniu wniosków.

Nie jest to jedyną możliwą formą rozumienia i interpretacji miary mocy efektu. Niezwykle konstruktywne jest działanie zmierzające do przekształcania miary mocy efektu na współczynniki korelacji. Do realizacji tego zamierzenia wymagane jest założenie mówiące, że kategorie grupy eksperymentalnej i kontrolnej należą do kategorii dychotomicznych zmiennych ciągłych. Roz-

Moc efektu w badaniach

woju procedury transformacji miary mocy efekty „d” na współczynnik korelacji dokonał Hedges (1986, s. 77). Przedstawia ją następujący wzór:

$$r \cong \frac{d}{\sqrt{\frac{d^2 + (n_e + n_k - 2)}{\frac{n_e n_k}{n_e + n_k}}}}$$

Gdy liczebność prób (grupy eksperymentalnej i kontrolnej) w badaniach eksperymentalnych jest jednakowa, nale y zastosować wzór otrzymany w wyniku redukcji:

$$r \cong \frac{d}{\sqrt{\frac{d^2 + 4(n-1)}{n}}}$$

Lub prostrzy:

$$r \cong \frac{d}{\sqrt{d^2 + 4}}$$

Otrzymane w ten sposób współczynniki „r” mo na interpretować podobnie jak inne współczynniki korelacji. Dokonując mno enia razy 100% ($r^2 \times 100\%$), mo na otrzymać wartość procentową wskazującą na moc oddziaływania czynnika eksperymentalnego na wariancję zmiennej zale nej.

Tabela 24 prezentuje wartości „d”, współczynniki korelacji oraz wartości procentowe uzyskane w prowadzonych eksperymentach z liczebnością grup wahającą się między dziesięć a dwadzieścia.

Wyniki tabeli wskazują wyraźnie, e stosunkowo du ym wartościom „d” odpowiadają niewielkie wartości procentowe wariancji.

Tabela 25 wskazuje na stosunek między typowymi wartościami „d” a procentem wariancji. We wszystkich eksperymentach liczebność grupy eksperymentalnej i grupy kontrolnej jest jednakowa (dla przypadków A liczebność wynosi $n = 8$, dla przypadków B liczebność $n = 25$). W eksperymentach dydaktycznych liczebność grup niezwykle rzadko jest mniejsza od 8 lub większa od 25 respondentów. Tabela ta umo liwia otrzymanie odpowiedzi na pytanie: czego nale y oczekiwać od prowadzonych metaanaliz badań podobnych do omówionych?

Zaprezentowany rozdział dotyczył do tej pory miary mocy efektu w badaniach eksperymentalnych. Przesłanką rozwa ań stał się szereg eksperymentów

Moc efektu w badaniach

Tabela 24. Wartości „d”, współczynniki korelacji oraz wartości procentowe uzyskane w 21 eksperymentach

| Badanie | Wartość „d” | Współczynnik korelacji | % |
|---------|-------------|------------------------|------|
| 1 | 0,79 | 0,38 | 14,8 |
| 2 | 0,38 | 0,19 | 3,8 |
| 3 | 0,55 | 0,27 | 7,4 |
| 4 | 0,71 | 0,34 | 11,8 |
| 5 | 0,21 | 0,11 | 1,2 |
| 6 | 0,33 | 0,17 | 2,9 |
| 7 | 0,67 | 0,32 | 10,6 |
| 8 | 0,45 | 0,23 | 5,3 |
| 9 | 0,52 | 0,26 | 7,0 |
| 10 | 0,48 | 0,25 | 6,0 |
| 11 | 0,86 | 0,41 | 16,8 |
| 12 | 0,85 | 0,41 | 16,7 |
| 13 | 0,66 | 0,33 | 10,8 |
| 14 | 0,73 | 0,36 | 12,9 |
| 15 | 1,01 | 0,46 | 21,4 |
| 16 | 0,62 | 0,31 | 9,5 |
| 17 | 0,44 | 0,22 | 4,9 |
| 18 | 0,56 | 0,28 | 8,0 |
| 19 | 1,21 | 0,53 | 27,9 |
| 20 | 0,28 | 0,15 | 2,1 |
| 21 | 1,12 | 0,50 | 24,9 |

ródło: badania własne.

z wyodrębnionym problemem badawczym, który był identyczny lub bardzo zbliżony. W każdym eksperymencie badaniom poddawano przynajmniej dwie grupy porównywalne. Z poszczególnych eksperymentów pobrano średnie arytmetyczne obydwu grup, wariancję grupy kontrolnej (lub wariancję wewnątrzgrupową) i dane zawierające liczebność prób. Są to wystarczające dane do wyliczenia miary mocy efektu. Pytania badawcze w metaanalizie wymagają czasami komparacji więcej niż dwóch kategorii czynnika eksperymentalnego. Są to sytuacje, w których występują więcej niż dwie grupy porównywalne. W takich przypadkach dokonuje się odpowiednio więcej odrębnych analiz statystycznych. Każdemu pytaniu badawczemu metaanalizy przyporządkowana jest jedna para grup z każdego badania pierwotnego. Rezultaty każdej pary są podstawą obliczeń mocy efektu. Analogiczna procedura obowiązuje przy poszukiwaniu wyników, stanowiących podstawę odpowiedzi na kolejne pytania badawcze. Nadmieniona odrębność czynności w metaanalizie

Tabela 25. Stosunek między typowymi wartościami „d” a wartościami procentowymi wariancji

| Wartość „d” | %(A) | %(B) |
|-------------|-------|-------|
| 0,00 | 0,00 | 0,00 |
| 0,05 | 0,07 | 0,07 |
| 0,10 | 0,29 | 0,26 |
| 0,15 | 0,64 | 0,58 |
| 0,20 | 1,13 | 1,03 |
| 0,25 | 1,75 | 1,60 |
| 0,30 | 2,51 | 2,29 |
| 0,35 | 3,38 | 3,09 |
| 0,40 | 4,37 | 4,00 |
| 0,45 | 5,47 | 5,01 |
| 0,50 | 6,67 | 6,11 |
| 0,60 | 9,33 | 8,57 |
| 0,70 | 12,28 | 11,32 |
| 0,80 | 15,46 | 14,29 |
| 0,90 | 18,79 | 17,42 |
| 1,00 | 22,22 | 20,66 |
| 1,10 | 25,69 | 23,96 |
| 1,20 | 29,15 | 27,27 |
| 1,30 | 32,56 | 30,56 |
| 1,40 | 35,90 | 33,91 |
| 1,50 | 39,13 | 36,95 |
| 2,00 | 53,33 | 51,02 |
| 3,00 | 72,00 | 70,09 |

ródło: badania własne.

dotyczy jedynie opracowań statystycznych, natomiast wnioskowanie ma charakter uogólniony, globalny i kompleksowy.

Większość badań empirycznych w obszarze pedagogiki nie należy do badań eksperymentalnych. W pedagogice istotniejsze od określania miary mocy efektu z eksperymentów jest obliczanie tej miary z badań nieeksperymentalnych.

Miara mocy efektu otrzymana w postaci standaryzowanej różnicy średnich arytmetycznych „d” została rozwinięta w celu integracji eksperymentów. Jednak nie zawiera ona żadnych cech, które ograniczałyby jej stosowanie tylko do metaanalizy eksperymentów. Z uwagi na swój charakter wartość „d” może być zastosowana przy dokonywaniu integracji wszystkich badań, w których rezultaty wyrażone zostały w postaci średnich arytmetycznych

Moc efektu w badaniach

oraz dokonano w nich porównania rezultatów przynajmniej dwóch grup. Zasadnicza różnica w stosowaniu miary „d” przy dokonywaniu integracji rezultatów badań eksperymentalnych i nieeksperymentalnych polega na realizacji procedury badawczej.

Badania nieeksperymentalne są bardziej zróżnicowane. Konsekwencją tego jest skomplikowany dobór badań do metaanalizy, kategoryzowanie cech wybranych do analizy statystycznej oraz opracowanie rezultatów. Przy wyliczaniu miary mocy efektu należy uwzględniać fakt polegający na tym, że w badaniach nieeksperymentalnych grupy poddane komparacji nie posiadają jednakowej liczby respondentów. W badaniach nieeksperymentalnych niezwykle rzadko stosuje się procedurę ujednoczenia grup. Fakt ten ma duży wpływ na rezultaty, dlatego też istnieje konieczność uwzględniania go w interpretacji wyników metaanalizy.

Wskazane przesłanki stanowią dowód, że opisana miara mocy efektu nie jest miarą odnoszącą się jedynie do eksperymentów. Eksperymenty przyjęte za punkt wyjścia w pierwszej fazie rozważań omawianego rozdziału, zostały omówione z uwagi na prostotę dokonywania ich integracji w metaanalizie. Podobieństwo problemów badawczych, zastosowanych procedur pomiaru i metod statystycznych oraz rezultatów końcowych jest większe w badaniach eksperymentalnych niż w przypadku badań nieeksperymentalnych.

Omówiony powyżej rozdział zawiera jedynie wybrane spośród wielu możliwości rozwiązań określania miary mocy efektu. Literatura metodologiczna określa je czasem w dużym uproszczeniu jako „miary podstawowe”. Są nimi objęte najczęściej występujące rezultaty empiryczne badań pierwotnych. Analiza omawianych rozwiązań wskazała na możliwości i ograniczenia rozwiniętych do dziś miar. Przedstawiono tu również główne dylematy, spotykane w praktyce metaanalizy przy stosowaniu miar mocy efektu oraz wpływ tych dylematów na procesy dalszego rozwoju i doskonalenia poszczególnych miar. Teoria i praktyka metaanalizy rozwijają z pewnością nowe podejścia, głównie w obszarze nieparametrycznych miar mocy efektu.

Określanie mocy efektu stanowi pierwszy krok w metaanalizy integracji rezultatów. Same miary mocy efektu nie wskazują na wszystkie oczekiwane informacje przez badacza. Problem badawczy metaanalizy wymaga dalszego opracowania miar mocy efektu. Stanowią one empiryczny materiał wyjściowy w metaanalizie, podobnie jak surowe (nieopracowane) dane – empiryczny materiał wyjściowy w badaniach pierwotnych.

Zakończenie

Zaprezentowane w tej różnorodnej koncepcje dotyczące przeprowadzenia i integracji badań pedagogicznych, stanowią zaledwie wycinek bogatej i obszernej teorii metodologicznej. Treści zawarte w poszczególnych rozdziałach i podrozdziałach koncentrują się wokół zasadniczych podejść metodologicznych.

Złożoność oraz różnorodność przedstawionych teorii coraz wyraźniej dowodzi tezy, iż dotychczasowe osiągnięcia powinny stać się jedynie podstawową bazą rozwoju dociekań naukowych. Niemal każda z zaprezentowanych w książce teorii i koncepcji nie jest wolna od rodzących się wątpliwości oraz konieczności dalszych poszukiwań.

W pracy zwrócono uwagę na rodzące się koncepcje wokół integracji ilościowych i jakościowych oraz związane z nimi polemiki. Wśród narosłych sporów i kontrowersji niepodważalne jest ustalenie, iż jednym z ważniejszych aspektów badań naukowych, przebiegających w miejscu przejścia od empirycznych do teoretycznych dociekań, jest problem integracji rezultatów badań empirycznych. Integracja ta stanowi pomost pomiędzy zdarzeniami specyficznymi i faktami już rozpoznanymi a prawami o charakterze bardziej ogólnym. Można przyjąć, iż próbę początkowej integracji stanowi już każde dobre pierwotne badanie empiryczne. Dokonana w ten sposób integracja jest zarówno jakościowa, jak i ilościowa. Historia integracji pokazuje wyraźnie, iż istniały różne tendencje w ustalaniu proporcji w jakościowej i ilościowej procedurze dociekań naukowych. Przykładem integracji głównie jakościowej stanowią znane od dawna w pedagogice przeglądy tradycyjne.

Rozwój metodologii nauk humanistycznych dał również szansę rozwinięcia badań z procedurą integracji o bardziej ilościowym charakterze. Sposób ten określono mianem metaanalizy. Powstanie bardziej złożonych i precyzyjnych procedur metaanalizy przypada na początek lat sześćdziesiątych dwudziestego wieku. Efektem ich rozwoju stały się publikacje pierwszych teoretycznych prac dotyczących idei, sensu i podstaw prowadzenia integracji badań. Okres ten to również czas rozkwitu wdrażania w praktykę funda-

Zakończenie

mentalnych zasad metodologicznych. W dokonywanych w tym czasie pierwszych metaanalizach badacze przedstawiali różne koncepcje integracji kwantytatywnej. Rozległość tych metaanaliz była zdecydowanie większa niż rozległość dotychczasowych przeglądów kwantytatywnych. Szybko stało się oczywiste, iż do przeprowadzania integracji są potrzebne oraz konieczne nowe metody statystyczne. Ich rozwój stał się możliwy głównie dzięki poszukiwaniom oraz rozważaniom Glassa, Smitha, Rosenthala i Hedgesa.

Najwcześniejsze metaanalizy dowodzą, że badacze dokonywali przeliczeń rezultatów w zbiorze integrowanych badań. Wraz z rozwojem badań rozwinięto te procedury badawcze. Nowo wprowadzoną czynnością stało się rozwarstwianie tych rezultatów najczęściej na trzy kategorie. Pierwszą z nich stanowiły rezultaty istotne w jednym kierunku, czyli wyniki potwierdzające przyjętą hipotezę. Na kategorię drugą składały się wyniki istotne w przeciwnym kierunku, czyli rezultaty potwierdzające hipotezę przeciwną. Trzecią, ostatnią kategorię stanowiły rezultaty nieistotne, czyli takie, które nie potwierdzały żadnej z przyjętych wcześniej hipotez.

Podstawą opracowania wniosków końcowych z tak dokonywanych integracji stała się wartość procentowa poszczególnych kategorii. Metoda ta posiadała jednak zasadniczą słabość – uwzględniała ona jedynie kierunek efektu poszczególnych badań, z pominięciem jego mocy.

Metaanalizy dokonywane później przyniosły rozwiązania polegające na rozważaniach wartości procentowej badanych zmian w kierunku określonej hipotezy. Możliwe stało się – w przypadku wartości procentowej – uwzględnianie nie tylko kierunku, ale również mocy efektu. Okazało się więc oczywiście, że niewystarczające jest już jedynie ustalenie, jak duża liczba wyników wartości procentowej potwierdza założoną hipotezę, ale istotne znaczenie ma też siła, z jaką określona wartość procentowa potwierdza założoną hipotezę.

Wadą omówionych czynności jest to, że większość z nich nie dokonuje jednak ścisłego pomiaru mocy efektu. W wielu przypadkach, jeżeli nawet wyniki te są uzyskiwane, to otrzymane miary mocy efektu są między sobą nieporównywalne. Staje się to powodem potencjalnych ograniczeń w procesie interpretacji wyników oraz podczas opracowywania wniosków końcowych. Problem ten wskazuje jednocześnie na konieczność odnalezienia metod umożliwiających określanie mocy efektu w jednakowej skali, która stanie się niezależna od jednostek pomiaru w poszczególnych badaniach. Postawione zadanie z punktu widzenia różnorodności rezultatów empirycznych, koniecznych do otrzymania miar mocy efektu, nie należy jednak do łatwych wyzwań. Różnorodność ta jest skutkiem istnienia wielu możliwości stosowania rozmaitych procedur pomiaru oraz różnych metod statystycznych przy opracowa-

Zakończenie

niu tego samego problemu badawczego. Stąd te za najbardziej obiecujące uznano postulaty rozwoju takich metod, które uwzględniają różnorodność rezultatów empirycznych badań pierwotnych. Często bowiem zdarza się, że różnorodność ta nie pozwala na włączenie wszystkich rezultatów do analizy statystycznej.

Wśród rezultatów badań empirycznych w obszarze pedagogiki pojawiają się najczęściej trzy parametry: średnie arytmetyczne, współczynniki korelacji oraz wartości procentowe.

Metaanaliza integruje przeważnie badania, które zawierają rezultaty wyrażone w postaci tego samego parametru. Zdecydowanie rzadziej integracji podlegają badania, których wyniki zostają określone za pomocą dwóch lub wszystkich trzech wymienionych wyżej parametrów.

Analizowane procedury stanowią narzędzie, dzięki któremu możliwe jest odnalezienie odpowiedzi na dwa zasadnicze pytania metaanalizy. Jedno z nich dotyczy określenia mocy efektu, stanowiącej wynik otrzymany z badań pojedynczych. Drugie z pytań zawiera problem wpływu poszczególnych cech badań na ich rezultaty. Głównym zamierzeniem autorów książki było ukazanie zależności dociekań teoretycznych i empirycznych w rozwoju metodologii pedagogicznej. Autorzy podjęli próbę zgłębienia podstawowych teorii, koncepcji i kontrowersji, które wpłynęły na burzliwy rozwój metodologii pedagogicznej i szczególnie tej subdyscypliny w obszarze nauk pedagogicznych, którą jest integracja badań.



Bibliografia

- Abbott J. A., Faris S. E. (2000), *Integrating technology into preservice literacy instructions*, „Journal of Research on Computing in Education”, 33, s. 149–161.
- Appenzeller O., Amm M., Jones H. (2004), *A brief exploration of neurological art history*, „Journal of the History of the Neurosciences”, 13, s. 345–350.
- Bangert-Drowns R. L., Kulik J. A., Kulik C.-L. C. (1983), *Effect of coaching programs on achievement test performance*, „Review of Educational Research”, s. 571–585.
- Becker H. J. (1988), *The impact of computer use on children’s learning*, Baltimore.
- Bloom B. S. (1976), *Human Characteristics and School Learning*, New York.
- Chu G. C., Schramm W. (1968), *Learning from television: What the research says*, Washington.
- Cohen J. (1988), *Statistical power analysis for the behavioral sciences*, Hillsdale.
- Cook T. D. (1992), *Some generic Issues and Problems for Meta-Analysis*, [w:] *Meta-analysis for explanation: A Casebook*, red. T. D. Cook, New York.
- Educational Research Service (1980), *Class size research: A critique of recent meta-analyses*, *Phi Delta Kappan*, s. 239–241.
- Erlenmeyer-Kimling L., Jarvik L. F. (1963), *Genetics and intelligence: A review*, „Science”, s. 1477–1479.
- Eysenck H. J. (1952), *The effects of psychotherapy: An evaluation*, „Journal of Consulting Psychology”, s. 319–324.
- Eysenck H. J. (1978), *An exercise in mega-silliness*, „American Psychologist”, s. 517.
- Fraser B. J. (1987), *Syntheses of Educational Productivity Research*, „International Journal of Educational Research”, s. 189–200.
- Glass G. V. (1976), *Primary, secondary, and meta-analysis of research*, „Educational Researcher”, s. 3–8.
- Glass G. V. (1980), *Summarizing effect sizes*, [w:] *New directions for methodology of social and behavioral science: Quantitative assessment of research domains*, red. R. Rosenthal, San Francisco.

Bibliografia

- Glass G. V., Smith M. L. (1979), *Meta-analysis of research on class size and achievement*, „Educational Evaluation and Policy Analysis”, s. 2–16.
- Glass G. V., Cahen L. S., Smith M. L., Filby N. N. (1982), *School class size: Research and policy*, Beverly Hills.
- Glass G. V., McGaw B., Smith M. L. (1981), *Meta-analysis in social research*, Beverly Hills.
- Hedges L. V. (1981), *Distribution theory for Glass's estimator of effect size and related estimators*, „Journal of Educational Statistics”, s. 107–128.
- Hedges L. V. (1984), *Advances in statistical methods for meta-analysis*, [w:] *Issues in data synthesis*, red. W. H. Yeaton i P. M. Wortman, San Francisco.
- Hedges L. V. (1986), *Issues in meta-analysis*, [w:] *Review of research in education*, red. E. Z. Rothkopf, Washington.
- Hedges L. V., Olkin I. (1985), *Statistical methods for meta-analysis*, New York.
- Hunter J. E., Schmidt F. L. (1990), *Methods of meta-analysis: Correcting error and bias in research findings*, Newbury Park.
- Hunter J. E., Schmidt F. L., Coggin T. D. (1988), *Meta-analysis of correlations: The issue of bias and misconceptions about the Fisher z transformation*, Michigan.
- Hunter J. E., Schmidt F. L., Jackson G. B. (1982), *Meta-analysis: Cumulating research findings across studies*, Beverly Hills.
- Kozłowska A. (1999), *Ocenianie w szkole na progach edukacyjnych*, Częstochowa.
- Kozłowska A. (2001), *Nie bój się pomiaru dydaktycznego w szkole czyli o testach osiągnięć uczniów*, Częstochowa.
- Kozłowska A. (2002), *Pomiar dydaktyczny i ewaluacja w szkole*, Częstochowa.
- Kożuh B. (1995), *Metaanaliza w badaniach szkolno-reformatorskich*, [w:] *Edukacja w procesie przemian cywilizacyjnych i kulturowych*, Opole.
- Kożuh B. (1999), *Metaanaliza*, Uniwersytet Opolski, Instytut Nauk Pedagogicznych, Opole.
- Kożuh B. (2005), *Metaanalitična integracija pedagoških raziskav*, Ljubljana.
- Kulik C.-L. C., Kulik J. A. (1982), *Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings*, „American Educational Research Journal”, s. 415–428.
- Kulik C.-L. C., Kulik J. A. (1985), *Estimating effect sizes in quantitative research integration*, Ann Arbor.
- Kulik C.-L. C., Kulik J. A. (1986), *Effectiveness of computer-based education in colleges*, „AEDS Journal”, s. 81–108.
- Kulik J. A., Kulik C.-L. C. (1987), *Effects of ability grouping on student achievement*, „Equity and Excellence”, s. 22–30.
- Kulik J. A., Kulik C.-L. C. (1989), *Meta-Analysis in Education*, „International Review of Educational Research”, s. 223–340.

Bibliografia

- Lipsey M. W. (1992), *Juvenile Delinquency Treatment: A Meta-Analytic Inquiry into the Variability of Effects*, [w:] *Meta-analysis for explanation: A Casebook*, New York.
- Mansfield R. S., Buse T. V. (1997), *Meta-analysis of research: A rejoinder to Glass*, „Educational Researcher”, s. 3.
- McGaw B. (1988), *Meta-analysis*, [w:] *Educational research, methodology and measurement: an international handbook*, red. J. P. Keeves, Oxford.
- Meehl P. E. (1954), *Clinical versus statistical prediction*, Minneapolis.
- Nowaczyk Cz. (1985), *Podstawy metod statystycznych dla pedagogów*, Warszawa–Poznań.
- Pilch T. (1995), *Zasady badań pedagogicznych*, Warszawa.
- Presby S. (1978), *Overly broad categories obscure important differences between therapies*, „American Psychologist”, s. 514–515.
- Rosenthal R. (1963), *On the social psychology of the psychological experiment*, „American Scientist”, s. 268–283.
- Rosenthal R. (1976), *Experimenter effects in behavioral research*, New York.
- Rosenthal R. (1990), *Evaluation of procedures and results*, [w:] *The future of meta-analysis*, red. K. Wachter, M. L. Straf New York.
- Rosenthal R. (1991), *Meta-analytic procedures for social research*, Beverly Hills.
- Sagadin J. (1993), *Poglavja iz metodologije pedagoškega raziskovanja*, Ljubljana.
- Slavin R. E. (1984), *Meta-analysis in education: How has it been used?*, „Educational Researcher”, s. 6–15.
- Slavin R. E. (1986), *Best-evidence synthesis: An alternative to meta-analysis and traditional reviews*, „Educational Researcher”, s. 5–11.
- Slavin R. E. (1987), *Ability grouping and student achievement in elementary schools: Best evidence synthesis*, „Review of Educational Research”, s. 293–336. (a)
- Slavin R. E. (1987), *Mastery learning reconsidered*, „Review of Educational Research”, s. 175–213. (b)
- Slavin R. E., Karweit N. (1984), *Within-class ability grouping and student achievement*, New Orleans.
- Smith M. L., Glass G. V., Miller T. I. (1980), *The benefits of psychotherapy*, Baltimore.
- Wolf F. M., Savickas M. L., Saltzman G. A., Walker M. L. (1984), *Meta-analytic evaluation of an interpersonal skills curriculum for medical students: Synthesizing evidence over successive occasions*, „Journal of Counseling Psychology”, s. 253–257.
- Zaczyński W. P. (1997), *Statystyka w pracy badawczej nauczyciela*, Warszawa.
- Бабанский Ю. К., *Проблемы повышения эффективности педагогических исследований: (Дидактический аспект)*, Москва: Педагогика, 1982.
- Барсков А. Г., *Научный метод: возможности и иллюзии*, Москва, 1994.

Bibliografia

Ботвинников А. Д., *Организация и методика педагогических исследований*, Москва: Наука, 1981.

Введение в научное исследование по педагогике: Учебное пособие для студентов педагогических институтов, Ю. К. Бабанский, В. И. Журавлев, В. К. Розов и др., Москва: Просвещение, 1988.

Георгиевский А. С., *Методология и методика научно-исследовательской работы*, Москва: Наука, 1982.

Глазиріна В. М., *Педагогіка сучасної школи: Навчальний посібник для студентів педагогічних ВНЗ*, Донецьк: Норд-Прес, 2006.

Давыдов В. П., *Основы методологии, методики и технологии педагогического исследования: Научно-методическое пособие*, Москва: Академия ФСБ, 1997.

Загвязинский В. И., Атаханов Р., *Методология и методы психолого-педагогического исследования: Учебное пособие*, Москва: Издательский центр „Академия”, 2001.

Зеленогорский Ф. А., *О методах исследования и доказательства*, Москва 1998.

Кохановский В. П., *Философия и методология науки*, Москва 1999.

Краевский В. В., *Методология педагогического исследования: Пособие для педагога-исследователя*, Самара: Изд-во СамГПИ, 1994.

Краевский В. В., *Соотношение педагогической науки и педагогической практики*, Москва 1977.

Кузьмина Н. В., *Методы системного педагогического исследования*, Ленинград 1982.

Найн А. Я., *Методология и методика научного исследования*, Челябинск, 1993.

Оценка качества научно-педагогических исследований, Москва 1987.

Полонский В. М., *Оценка качества научно-педагогических исследований*, Москва 1987.

Приходько П. Т., *Азбука исследовательского труда*, Москва: Наука, 1979.

Скалкова Я., *Методология, методы педагогического исследования*, пер. с чешск., Москва 1989.

Скаткин М. Н., *Методология и методика педагогических исследований (В помощь начинающему исследователю)*, Москва 1986.

Шевандрин Н. И., *Социальная психология в образовании: Учебное пособие. Ч. 1: Концептуальные и прикладные основы социальной психологии*, Москва: ВЛАДОС, 1995.

Aneksy

Tabela A. Wpływ niektórych czynników na osiągnięcia szkolne uczniów

| Niezależne czynniki i zmienne | % rozrzutu |
|---|------------|
| A. Wiadomości wstępne uczniów | 50 |
| B. Motywacja początkowa uczniów | 25 |
| a) motywacja ukierunkowana na poszczególne przedmioty | 12–20 |
| b) motywacja ukierunkowana na naukę i szkołę | 20 |
| c) motywacja ukierunkowana na samoocenę | 25 |
| C. Jakość nauczania | 25 |
| a) cechy nauczycieli | 5 |
| b) cechy klas i szkół | 5 |

ródło: Bloom 1976.

Tabela B. Rezultaty badania zależności między wybranymi czynnikami a rezultatami kształcenia przeprowadzonego przez Frasera

| Czynniki | Liczba opracowanych metaanaliz | Współczynnik korelacji | |
|-----------------------------|--------------------------------|------------------------|---------------------|
| | | dla całej grupy | maksymalny w grupie |
| Szkoła | 16 | 0,12 | 0,26 |
| Nauczyciel | 9 | 0,21 | 0,20 |
| Uczeń | 25 | 0,24 | 0,44 |
| Środowisko rodzinne ucznia | 4 | 0,19 | 0,31 |
| Organizacja procesu uczenia | 31 | 0,22 | 0,47 |
| Metody nauczania | 37 | 0,14 | 0,25 |
| Strategie nauczania | 12 | 0,28 | 0,49 |

ródło: Fraser 1987.

Aneksy

Tabela C. Powierzchnia pod krzywą normalną (%) określona odchyleniem od średniej mierzonym w jednostkach odchylenia standardowego (z)

| z | % | z | % | z | % | z | % | z | % |
|------|-------|------|-------|------|-------|------|-------|------|-------|
| 0,00 | 0,00 | 0,39 | 15,17 | 0,78 | 28,23 | 1,17 | 37,90 | 1,56 | 44,06 |
| 0,01 | 0,40 | 0,40 | 15,54 | 0,79 | 28,52 | 1,18 | 38,10 | 1,57 | 44,18 |
| 0,02 | 0,80 | 0,41 | 15,91 | 0,80 | 28,81 | 1,19 | 38,30 | 1,58 | 44,29 |
| 0,03 | 1,20 | 0,42 | 16,28 | 0,81 | 29,10 | 1,20 | 38,49 | 1,59 | 44,41 |
| 0,04 | 1,60 | 0,43 | 16,64 | 0,82 | 29,39 | 1,21 | 38,69 | 1,60 | 44,52 |
| 0,05 | 1,99 | 0,44 | 17,00 | 0,83 | 29,67 | 1,22 | 38,83 | 1,61 | 44,63 |
| 0,06 | 2,39 | 0,45 | 17,36 | 0,84 | 29,95 | 1,23 | 39,07 | 1,62 | 44,74 |
| 0,07 | 2,79 | 0,46 | 17,72 | 0,85 | 30,23 | 1,24 | 39,25 | 1,63 | 44,84 |
| 0,08 | 3,19 | 0,47 | 18,08 | 0,86 | 30,51 | 1,25 | 39,44 | 1,64 | 44,95 |
| 0,09 | 3,59 | 0,48 | 18,44 | 0,87 | 30,78 | 1,26 | 39,62 | 1,65 | 45,05 |
| 0,10 | 3,98 | 0,49 | 18,79 | 0,88 | 31,06 | 1,27 | 39,80 | 1,66 | 45,15 |
| 0,11 | 4,38 | 0,50 | 19,15 | 0,89 | 31,33 | 1,28 | 39,97 | 1,67 | 45,25 |
| 0,12 | 4,78 | 0,51 | 19,50 | 0,90 | 31,59 | 1,29 | 40,15 | 1,68 | 45,35 |
| 0,13 | 5,17 | 0,52 | 19,85 | 0,91 | 31,86 | 1,30 | 40,32 | 1,69 | 45,45 |
| 0,14 | 5,57 | 0,53 | 20,19 | 0,92 | 32,12 | 1,31 | 40,49 | 1,70 | 45,54 |
| 0,15 | 5,96 | 0,54 | 20,54 | 0,93 | 32,38 | 1,32 | 40,66 | 1,71 | 45,64 |
| 0,16 | 6,36 | 0,55 | 20,88 | 0,94 | 32,64 | 1,33 | 40,82 | 1,72 | 45,73 |
| 0,17 | 6,75 | 0,56 | 21,23 | 0,95 | 32,89 | 1,34 | 40,99 | 1,73 | 45,82 |
| 0,18 | 7,14 | 0,57 | 21,57 | 0,96 | 33,15 | 1,35 | 41,15 | 1,74 | 45,91 |
| 0,19 | 7,53 | 0,58 | 21,90 | 0,97 | 33,40 | 1,36 | 41,31 | 1,75 | 45,99 |
| 0,20 | 7,93 | 0,59 | 22,24 | 0,98 | 33,65 | 1,37 | 41,47 | 1,76 | 46,08 |
| 0,21 | 8,32 | 0,60 | 22,57 | 0,99 | 33,89 | 1,38 | 41,62 | 1,77 | 46,16 |
| 0,22 | 8,71 | 0,61 | 22,91 | 1,00 | 34,13 | 1,39 | 41,77 | 1,78 | 46,25 |
| 0,23 | 9,10 | 0,62 | 23,24 | 1,01 | 34,39 | 1,40 | 41,92 | 1,79 | 46,33 |
| 0,24 | 9,48 | 0,63 | 23,57 | 1,02 | 34,61 | 1,41 | 42,07 | 1,80 | 46,41 |
| 0,25 | 9,87 | 0,64 | 23,89 | 1,03 | 34,85 | 1,42 | 42,22 | 1,81 | 46,49 |
| 0,26 | 10,26 | 0,65 | 24,22 | 1,04 | 35,08 | 1,43 | 42,36 | 1,82 | 46,56 |
| 0,27 | 10,64 | 0,66 | 24,54 | 1,05 | 35,31 | 1,44 | 42,51 | 1,83 | 46,64 |
| 0,28 | 11,03 | 0,67 | 24,86 | 1,06 | 35,54 | 1,45 | 42,65 | 1,84 | 46,71 |
| 0,29 | 11,41 | 0,68 | 25,17 | 1,07 | 35,77 | 1,46 | 42,79 | 1,85 | 46,78 |
| 0,30 | 11,79 | 0,69 | 25,49 | 1,08 | 35,99 | 1,47 | 42,92 | 1,86 | 46,86 |
| 0,31 | 12,17 | 0,70 | 25,80 | 1,09 | 36,21 | 1,48 | 43,06 | 1,87 | 46,93 |
| 0,32 | 12,55 | 0,71 | 26,11 | 1,10 | 36,43 | 1,49 | 43,19 | 1,88 | 46,99 |
| 0,33 | 12,93 | 0,72 | 26,42 | 1,11 | 36,65 | 1,50 | 43,32 | 1,89 | 47,06 |
| 0,34 | 13,31 | 0,73 | 26,73 | 1,12 | 36,86 | 1,51 | 43,45 | 1,90 | 47,13 |
| 0,35 | 13,68 | 0,74 | 27,04 | 1,13 | 37,08 | 1,52 | 43,57 | 1,91 | 47,19 |
| 0,36 | 14,06 | 0,75 | 27,34 | 1,14 | 37,29 | 1,53 | 43,70 | 1,92 | 47,26 |
| 0,37 | 14,43 | 0,76 | 27,64 | 1,15 | 37,49 | 1,54 | 43,82 | 1,93 | 47,32 |
| 0,38 | 14,80 | 0,77 | 27,94 | 1,16 | 37,70 | 1,55 | 43,94 | 1,94 | 47,38 |

Aneksy

| z | % | z | % | z | % | z | % | z | % |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1,95 | 47,44 | 2,28 | 48,87 | 2,61 | 49,55 | 2,94 | 49,84 | 3,26 | 49,94 |
| 1,96 | 47,50 | 2,29 | 48,90 | 2,62 | 49,56 | 2,95 | 49,84 | 3,27 | 49,95 |
| 1,97 | 47,56 | 2,30 | 48,93 | 2,63 | 49,57 | 2,96 | 49,85 | 3,28 | 49,95 |
| 1,98 | 47,61 | 2,31 | 48,96 | 2,64 | 49,59 | 2,97 | 49,85 | 3,29 | 49,95 |
| 1,99 | 47,67 | 2,32 | 48,98 | 2,65 | 49,60 | 2,98 | 49,86 | 3,30 | 49,95 |
| 2,00 | 47,72 | 2,33 | 49,01 | 2,66 | 49,61 | 2,99 | 49,86 | 3,31 | 49,95 |
| 2,01 | 47,78 | 2,34 | 49,04 | 2,67 | 49,62 | 3,00 | 49,87 | 3,32 | 49,95 |
| 2,02 | 47,83 | 2,35 | 49,06 | 2,68 | 49,63 | 3,01 | 49,87 | 3,33 | 49,96 |
| 2,03 | 47,88 | 2,36 | 49,09 | 2,69 | 49,64 | 3,02 | 49,87 | 3,34 | 49,96 |
| 2,04 | 47,93 | 2,37 | 49,11 | 2,70 | 49,65 | 3,03 | 49,88 | 3,35 | 49,96 |
| 2,05 | 47,98 | 2,38 | 49,13 | 2,71 | 49,66 | 3,04 | 49,88 | 3,36 | 49,96 |
| 2,06 | 48,03 | 2,39 | 49,16 | 2,72 | 49,67 | 3,05 | 49,89 | 3,37 | 49,96 |
| 2,07 | 48,08 | 2,40 | 49,18 | 2,73 | 49,68 | 3,06 | 49,89 | 3,38 | 49,96 |
| 2,08 | 48,12 | 2,41 | 49,20 | 2,74 | 49,69 | 3,07 | 49,89 | 3,39 | 49,97 |
| 2,09 | 48,17 | 2,42 | 49,22 | 2,75 | 49,70 | 3,08 | 49,90 | 3,40 | 49,97 |
| 2,10 | 48,21 | 2,43 | 49,25 | 2,76 | 49,71 | 3,09 | 49,90 | 3,41 | 49,97 |
| 2,11 | 48,26 | 2,44 | 49,27 | 2,77 | 49,72 | 3,10 | 49,90 | 3,42 | 49,97 |
| 2,12 | 48,30 | 2,45 | 49,29 | 2,78 | 49,73 | 3,11 | 49,91 | 3,43 | 49,97 |
| 2,13 | 48,34 | 2,46 | 49,31 | 2,79 | 49,74 | 3,12 | 49,91 | 3,44 | 49,97 |
| 2,14 | 48,38 | 2,47 | 49,32 | 2,80 | 49,74 | 3,13 | 49,91 | 3,45 | 49,97 |
| 2,15 | 48,42 | 2,48 | 49,34 | 2,81 | 49,75 | 3,14 | 49,92 | 3,46 | 49,97 |
| 2,16 | 48,46 | 2,49 | 49,36 | 2,82 | 49,76 | 3,15 | 49,92 | 3,47 | 49,97 |
| 2,17 | 48,50 | 2,50 | 49,38 | 2,83 | 49,77 | 3,16 | 49,92 | 3,48 | 49,97 |
| 2,18 | 48,54 | 2,51 | 49,40 | 2,84 | 49,77 | 3,17 | 49,92 | 3,49 | 49,98 |
| 2,19 | 48,57 | 2,52 | 49,41 | 2,85 | 49,78 | 3,18 | 49,93 | 3,50 | 49,98 |
| 2,20 | 48,61 | 2,53 | 49,43 | 2,86 | 49,79 | 3,19 | 49,93 | 3,60 | 49,98 |
| 2,21 | 48,64 | 2,54 | 49,45 | 2,87 | 49,79 | 3,20 | 49,93 | 3,70 | 49,99 |
| 2,22 | 48,68 | 2,55 | 49,46 | 2,88 | 49,80 | 3,21 | 49,93 | 3,80 | 49,99 |
| 2,23 | 48,71 | 2,56 | 49,48 | 2,89 | 49,81 | 3,22 | 49,94 | 3,90 | 49,99 |
| 2,24 | 48,75 | 2,57 | 49,49 | 2,90 | 49,81 | 3,23 | 49,94 | 4,00 | 49,997 |
| 2,25 | 48,78 | 2,58 | 49,51 | 2,91 | 49,82 | 3,24 | 49,94 | 4,50 | 49,9998 |
| 2,26 | 48,81 | 2,59 | 49,52 | 2,92 | 49,82 | 3,25 | 49,94 | 5,00 | 49,99997 |
| 2,27 | 48,84 | 2,60 | 49,53 | 2,93 | 49,83 | | | | |

Aneksy

Tabela D. Korekcja Hedgesa

| $n_c + n_k - 2$ | J | $n_c + n_k - 2$ | J | $n_c + n_k - 2$ | J | $n_c + n_k - 2$ | J |
|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|
| 2 | 0,5642 | 15 | 0,9490 | 28 | 0,9729 | 41 | 0,9816 |
| 3 | 0,7236 | 16 | 0,9523 | 29 | 0,9739 | 42 | 0,9820 |
| 4 | 0,7979 | 17 | 0,9551 | 30 | 0,9748 | 43 | 0,9824 |
| 5 | 0,8048 | 18 | 0,9577 | 31 | 0,9756 | 44 | 0,9828 |
| 6 | 0,8686 | 19 | 0,9599 | 32 | 0,9764 | 45 | 0,9832 |
| 7 | 0,8882 | 20 | 0,9619 | 33 | 0,9771 | 46 | 0,9836 |
| 8 | 0,9027 | 21 | 0,9638 | 34 | 0,9778 | 47 | 0,9839 |
| 9 | 0,9139 | 22 | 0,9655 | 35 | 0,9784 | 48 | 0,9843 |
| 10 | 0,9228 | 23 | 0,9670 | 36 | 0,9790 | 49 | 0,9846 |
| 11 | 0,9300 | 24 | 0,9684 | 37 | 0,9796 | 50 | 0,9849 |
| 12 | 0,9359 | 25 | 0,9699 | 38 | 0,9801 | 51 | 0,9852 |
| 13 | 0,9410 | 26 | 0,9708 | 39 | 0,9806 | 52 | 0,9855 |
| 14 | 0,9453 | 27 | 0,9719 | 40 | 0,9811 | 53 | 0,9858 |

Aneksy

Tabela E. Zmiana r Pearsona na odpowiednie Z_r Fischera

| r | Z_r | r | Z_r | r | Z_r | r | Z_r |
|------|-------|------|-------|------|-------|------|-------|
| 0,00 | 0,000 | 0,25 | 0,255 | 0,50 | 0,549 | 0,75 | 0,973 |
| 0,01 | 0,010 | 0,26 | 0,266 | 0,51 | 0,563 | 0,76 | 0,996 |
| 0,02 | 0,020 | 0,27 | 0,277 | 0,52 | 0,576 | 0,77 | 1,020 |
| 0,03 | 0,030 | 0,28 | 0,288 | 0,53 | 0,590 | 0,78 | 1,045 |
| 0,04 | 0,040 | 0,29 | 0,299 | 0,54 | 0,604 | 0,79 | 1,071 |
| | | | | | | | |
| 0,05 | 0,050 | 0,30 | 0,310 | 0,55 | 0,618 | 0,80 | 1,099 |
| 0,06 | 0,060 | 0,31 | 0,321 | 0,56 | 0,633 | 0,81 | 1,127 |
| 0,07 | 0,070 | 0,32 | 0,332 | 0,57 | 0,648 | 0,82 | 1,157 |
| 0,08 | 0,080 | 0,33 | 0,343 | 0,58 | 0,662 | 0,83 | 1,188 |
| 0,09 | 0,090 | 0,34 | 0,354 | 0,59 | 0,678 | 0,84 | 1,221 |
| | | | | | | | |
| 0,10 | 0,100 | 0,35 | 0,365 | 0,60 | 0,693 | 0,85 | 1,256 |
| 0,11 | 0,110 | 0,36 | 0,377 | 0,61 | 0,709 | 0,86 | 1,293 |
| 0,12 | 0,121 | 0,37 | 0,388 | 0,62 | 0,725 | 0,87 | 1,333 |
| 0,13 | 0,131 | 0,38 | 0,400 | 0,63 | 0,741 | 0,88 | 1,376 |
| 0,14 | 0,141 | 0,39 | 0,412 | 0,64 | 0,758 | 0,89 | 1,422 |
| | | | | | | | |
| 0,15 | 0,151 | 0,40 | 0,424 | 0,65 | 0,775 | 0,90 | 1,472 |
| 0,16 | 0,161 | 0,41 | 0,436 | 0,66 | 0,793 | 0,91 | 1,528 |
| 0,17 | 0,172 | 0,42 | 0,448 | 0,67 | 0,811 | 0,92 | 1,589 |
| 0,18 | 0,182 | 0,43 | 0,460 | 0,68 | 0,829 | 0,93 | 1,658 |
| 0,19 | 0,192 | 0,44 | 0,472 | 0,69 | 0,848 | 0,94 | 1,738 |
| | | | | | | | |
| 0,20 | 0,203 | 0,45 | 0,485 | 0,70 | 0,867 | 0,95 | 1,832 |
| 0,21 | 0,213 | 0,46 | 0,497 | 0,71 | 0,887 | 0,96 | 1,946 |
| 0,22 | 0,224 | 0,47 | 0,510 | 0,72 | 0,908 | 0,97 | 2,092 |
| 0,23 | 0,234 | 0,48 | 0,523 | 0,73 | 0,929 | 0,98 | 2,298 |
| 0,24 | 0,245 | 0,49 | 0,536 | 0,74 | 0,950 | 0,99 | 2,647 |

Aneksy

Tabela F. Zmiana Z_r Fischera na odpowiednie r Pearsona

| Z_r | r | Z_r | r | Z_r | r | Z_r | r |
|-------|--------|-------|--------|-------|--------|-------|--------|
| 0,00 | 0,000 | 0,39 | 0,371 | 0,78 | 0,653 | 1,17 | 0,824 |
| 0,01 | 0,010 | 0,40 | 0,380 | 0,79 | 0,658 | 1,18 | 0,828 |
| 0,02 | 0,020 | 0,41 | 0,389 | 0,80 | 0,664 | 1,19 | 0,831 |
| 0,03 | 0,030 | 0,42 | 0,397 | 0,81 | 0,670 | 1,20 | 0,834 |
| 0,04 | 0,040 | 0,43 | 0,405 | 0,82 | 0,675 | 1,21 | 0,837 |
| 0,05 | 0,050 | 0,44 | 0,414 | 0,83 | 0,6805 | 1,22 | 0,840 |
| 0,06 | 0,060 | 0,45 | 0,422 | 0,84 | 0,686 | 1,23 | 0,843 |
| 0,07 | 0,070 | 0,46 | 0,430 | 0,85 | 0,691 | 1,24 | 0,846 |
| 0,08 | 0,080 | 0,47 | 0,438 | 0,86 | 0,696 | 1,25 | 0,848 |
| 0,09 | 0,090 | 0,48 | 0,446 | 0,87 | 0,701 | 1,26 | 0,851 |
| 0,10 | 0,100 | 0,49 | 0,454 | 0,88 | 0,706 | 1,27 | 0,854 |
| 0,11 | 0,110 | 0,50 | 0,462 | 0,89 | 0,711 | 1,28 | 0,8565 |
| 0,12 | 0,119 | 0,51 | 0,470 | 0,90 | 0,716 | 1,29 | 0,859 |
| 0,13 | 0,129 | 0,52 | 0,478 | 0,91 | 0,721 | 1,30 | 0,862 |
| 0,14 | 0,139 | 0,53 | 0,485 | 0,92 | 0,726 | 1,31 | 0,864 |
| 0,15 | 0,149 | 0,54 | 0,493 | 0,93 | 0,731 | 1,32 | 0,867 |
| 0,16 | 0,159 | 0,55 | 0,5005 | 0,94 | 0,735 | 1,33 | 0,869 |
| 0,17 | 0,168 | 0,56 | 0,508 | 0,95 | 0,740 | 1,34 | 0,872 |
| 0,18 | 0,178 | 0,57 | 0,515 | 0,96 | 0,744 | 1,35 | 0,874 |
| 0,19 | 0,188 | 0,58 | 0,523 | 0,97 | 0,749 | 1,36 | 0,876 |
| 0,20 | 0,197 | 0,59 | 0,530 | 0,98 | 0,753 | 1,37 | 0,879 |
| 0,21 | 0,207 | 0,60 | 0,537 | 0,99 | 0,757 | 1,38 | 0,881 |
| 0,22 | 0,2165 | 0,61 | 0,544 | 1,00 | 0,762 | 1,39 | 0,883 |
| 0,23 | 0,226 | 0,62 | 0,551 | 1,01 | 0,766 | 1,40 | 0,885 |
| 0,24 | 0,236 | 0,63 | 0,558 | 1,02 | 0,770 | 1,41 | 0,888 |
| 0,25 | 0,245 | 0,64 | 0,565 | 1,03 | 0,774 | 1,42 | 0,890 |
| 0,26 | 0,254 | 0,65 | 0,572 | 1,04 | 0,778 | 1,43 | 0,892 |
| 0,27 | 0,264 | 0,66 | 0,578 | 1,05 | 0,782 | 1,44 | 0,894 |
| 0,28 | 0,273 | 0,67 | 0,585 | 1,06 | 0,786 | 1,45 | 0,896 |
| 0,29 | 0,282 | 0,68 | 0,592 | 1,07 | 0,790 | 1,46 | 0,898 |
| 0,30 | 0,291 | 0,69 | 0,598 | 1,08 | 0,793 | 1,47 | 0,900 |
| 0,31 | 0,300 | 0,70 | 0,604 | 1,09 | 0,797 | 1,48 | 0,902 |
| 0,32 | 0,310 | 0,71 | 0,611 | 1,10 | 0,8005 | 1,49 | 0,903 |
| 0,33 | 0,319 | 0,72 | 0,617 | 1,11 | 0,804 | 1,50 | 0,905 |
| 0,34 | 0,327 | 0,73 | 0,623 | 1,12 | 0,808 | 1,51 | 0,907 |
| 0,35 | 0,336 | 0,74 | 0,629 | 1,13 | 0,811 | 1,52 | 0,909 |
| 0,36 | 0,345 | 0,75 | 0,635 | 1,14 | 0,814 | 1,53 | 0,910 |
| 0,37 | 0,354 | 0,76 | 0,641 | 1,15 | 0,818 | 1,54 | 0,912 |
| 0,38 | 0,363 | 0,77 | 0,647 | 1,16 | 0,821 | 1,55 | 0,914 |

Aneksy

| Z_r | r | Z_r | r | Z_r | r | Z_r | r |
|-------|--------|-------|-------|-------|-------|-------|-------|
| 1,56 | 0,915 | 1,92 | 0,958 | 2,28 | 0,979 | 2,64 | 0,990 |
| 1,57 | 0,917 | 1,93 | 0,959 | 2,29 | 0,980 | 2,65 | 0,990 |
| 1,58 | 0,919 | 1,94 | 0,960 | 2,30 | 0,980 | 2,66 | 0,990 |
| 1,59 | 0,920 | 1,95 | 0,960 | 2,31 | 0,980 | 2,67 | 0,990 |
| 1,60 | 0,922 | 1,96 | 0,961 | 2,32 | 0,981 | 2,68 | 0,991 |
| 1,61 | 0,923 | 1,97 | 0,962 | 2,33 | 0,981 | 2,69 | 0,991 |
| 1,62 | 0,925 | 1,98 | 0,963 | 2,34 | 0,982 | 2,70 | 0,991 |
| 1,63 | 0,926 | 1,99 | 0,963 | 2,35 | 0,982 | 2,71 | 0,991 |
| 1,64 | 0,928 | 2,00 | 0,964 | 2,36 | 0,982 | 2,72 | 0,991 |
| 1,65 | 0,929 | 2,01 | 0,965 | 2,37 | 0,983 | 2,73 | 0,992 |
| 1,66 | 0,930 | 2,02 | 0,965 | 2,38 | 0,983 | 2,74 | 0,992 |
| 1,67 | 0,932 | 2,03 | 0,966 | 2,39 | 0,983 | 2,75 | 0,992 |
| 1,68 | 0,933 | 2,04 | 0,967 | 2,40 | 0,984 | 2,76 | 0,992 |
| 1,69 | 0,934 | 2,05 | 0,967 | 2,41 | 0,984 | 2,77 | 0,992 |
| 1,70 | 0,935 | 2,06 | 0,968 | 2,42 | 0,984 | 2,78 | 0,992 |
| 1,71 | 0,937 | 2,07 | 0,969 | 2,43 | 0,985 | 2,79 | 0,992 |
| 1,72 | 0,938 | 2,08 | 0,969 | 2,44 | 0,985 | 2,80 | 0,993 |
| 1,73 | 0,939 | 2,09 | 0,970 | 2,45 | 0,985 | 2,81 | 0,993 |
| 1,74 | 0,940 | 2,10 | 0,970 | 2,46 | 0,986 | 2,82 | 0,993 |
| 1,75 | 0,941 | 2,11 | 0,971 | 2,47 | 0,986 | 2,83 | 0,993 |
| 1,76 | 0,9425 | 2,12 | 0,972 | 2,48 | 0,986 | 2,84 | 0,993 |
| 1,77 | 0,944 | 2,13 | 0,972 | 2,49 | 0,986 | 2,85 | 0,993 |
| 1,78 | 0,945 | 2,14 | 0,973 | 2,50 | 0,987 | 2,86 | 0,993 |
| 1,79 | 0,946 | 2,15 | 0,973 | 2,51 | 0,987 | 2,87 | 0,994 |
| 1,80 | 0,947 | 2,16 | 0,974 | 2,52 | 0,987 | 2,88 | 0,994 |
| 1,81 | 0,948 | 2,17 | 0,974 | 2,53 | 0,987 | 2,89 | 0,994 |
| 1,82 | 0,949 | 2,18 | 0,975 | 2,54 | 0,988 | 2,90 | 0,994 |
| 1,83 | 0,950 | 2,19 | 0,975 | 2,55 | 0,988 | 2,91 | 0,994 |
| 1,84 | 0,951 | 2,20 | 0,976 | 2,56 | 0,988 | 2,92 | 0,994 |
| 1,85 | 0,952 | 2,21 | 0,976 | 2,57 | 0,988 | 2,93 | 0,994 |
| 1,86 | 0,953 | 2,22 | 0,977 | 2,58 | 0,989 | 2,94 | 0,994 |
| 1,87 | 0,954 | 2,23 | 0,977 | 2,59 | 0,989 | 2,95 | 0,995 |
| 1,88 | 0,954 | 2,24 | 0,978 | 2,60 | 0,989 | 2,96 | 0,995 |
| 1,89 | 0,955 | 2,25 | 0,978 | 2,61 | 0,989 | 2,97 | 0,995 |
| 1,90 | 0,956 | 2,26 | 0,978 | 2,62 | 0,989 | 2,98 | 0,995 |
| 1,91 | 0,957 | 2,27 | 0,979 | 2,63 | 0,990 | 2,99 | 0,995 |



Spis treści

| | |
|---|----|
| Wstęp | 5 |
| Rozdział I. Podstawy metodologii pedagogiki | 7 |
| 1.1. Wstęp | 7 |
| 1.1.1. Metody i środki badań | 16 |
| 1.2. Metodologiczne zasady badań | 17 |
| 1.3. Formy działalności poznawczej w dziedzinie pedagogiki | 24 |
| 1.3.1. Nauka jako instytucja społeczna | 28 |
| 1.3.2. Cechy indywidualnej działalności naukowej | 29 |
| 1.4. Metoda poznania naukowego | 30 |
| 1.4.1. Klasyfikacja metod poznania naukowego | 32 |
| 1.4.2. Klasyfikacja metod badań psychologiczno-pedagogicznych | 35 |
| 1.4.3. Ogólnonaukowe logiczne metody i sposoby poznania | 37 |
| Rozdział II. Metody badań pedagogicznych | 47 |
| 2.1. Badania empiryczne w pedagogice | 48 |
| 2.2. Badania teoretyczne w pedagogice | 55 |
| 2.3. Statystyczno-matematyczne metody badań pedagogicznych | 57 |
| 2.4. Zintegrowane metody badań pedagogicznych | 59 |
| Rozdział III. Miejsce metaanalizy w badaniach pedagogicznych | 63 |
| 3.1. Rozwój badań metaanalizy | 64 |
| 3.2. Główne cechy metaanalizy | 67 |
| 3.3. Koncepcje rozwoju metaanalizy | 69 |
| 3.3.1. Badania Rosenthala | 70 |
| 3.3.2. Restrykcja wyboru badań | 72 |
| 3.3.3. Ocenianie mocy efektu | 73 |
| 3.3.4. Integracja metaanaliz | 74 |
| 3.3.5. Krytyka metaanalizy | 76 |
| Rozdział IV. ródła w integracji badań pedagogicznych | 79 |
| 4.1. Definiowanie problemu badawczego | 79 |
| 4.2. Kryteria doboru badań | 81 |

Spis treści

| | |
|---|-----|
| 4.3. Procedury poszukiwania badań | 84 |
| 4.4. Przesiew wyników badań | 87 |
| | |
| Rozdział V. Opis cech badań | 91 |
| 5.1. Opis cech tematycznych | 92 |
| 5.2. Opis cech metodologicznych | 93 |
| | |
| Rozdział VI. Kategoryzacja badań | 101 |
| 6.1. Trafność | 101 |
| 6.2. Rzetelność | 102 |
| 6.2.1. Wartości procentowe jako wskaźniki rzetelności | 103 |
| 6.2.2. Współczynniki zbite ności i korelacji jako wskaźniki rzetelności | 108 |
| 6.2.2.1. Współczynnik zbite ności Pearsona | 108 |
| 6.2.2.2. Współczynnik zbite ności Cramera | 112 |
| 6.2.2.3. Współczynnik zbite ności ϕ | 113 |
| 6.2.2.4. Współczynnik korelacji Pearsona | 114 |
| 6.2.2.5. Stosunek korelacji | 115 |
| 6.2.2.6. Tetrachoryczny współczynnik korelacji | 116 |
| 6.2.2.7. Określanie rzetelności metodą reprezentatywną z rezultatów prób | 116 |
| 6.2.2.7.1. Ocenianie procentu badań jednakowo skategoryzowanych z próby | 117 |
| 6.2.2.7.2. Ocenianie współczynnika korelacji Pearsona z próby | 117 |
| | |
| Rozdział VII. Moc efektu w badaniach | 119 |
| 7.1. Miara mocy efektu średnich arytmetycznych | 122 |
| 7.2. Miara mocy efektu współczynników korelacji | 135 |
| 7.2.1. Współczynnik korelacji Pearsona | 135 |
| 7.2.2. Współczynnik Fishera | 136 |
| 7.2.3. Miara mocy efektu Rosenthala | 137 |
| 7.2.4. Pozostałe współczynniki korelacji | 138 |
| 7.2.5. Różnica współczynników korelacji | 138 |
| 7.3. Miara mocy efektu wartości procentowych | 139 |
| 7.4. Interpretacja mocy efektu | 141 |
| | |
| Zakończenie | 147 |
| Bibliografia | 151 |
| Aneksy | 155 |